

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2002, 08:48:45 ; Search time 40 Seconds  
(without alignments)

2368.530 Million cell updates/sec

Title: US-09-966-561-2

Perfect score: 3792

Sequence: 1 MAERESGLGGGAASPPAAS.....QQFYKQFVEYTCPTEDIYLE 711

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

1: /SID52/qcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID52/qcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID52/qcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID52/qcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID52/qcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID52/qcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID52/qcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID52/qcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID52/qcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID52/qcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID52/qcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID52/qcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID52/qcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID52/qcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID52/qcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID52/qcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID52/qcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID52/qcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID52/qcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID52/qcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID52/qcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID52/qcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID52/qcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3780	99.7	711	19	AAW80602 Human transcriptio
2	3544.5	93.5	708	22	AAW80602 Rat SAPK interacti
3	3544.5	93.5	715	22	AAW80602 Rat SAPK interacti
4	3529	93.1	707	22	AAW80602 Murine SKIP-1b. M
5	3503	92.4	714	19	AAW80601 Rat transcription
6	3389	89.4	699	22	AAW80602 Rat SAPK interacti
7	3367.5	88.8	698	22	AAW80602 Murine JNK-interac
8	3262.5	86.0	660	20	AAW81525 Murine JNK-interac
9	3248	85.7	673	22	AAW81525 Murine SKIP-2b. M
10	3245	85.6	659	23	ABB04788 LDL receptor bindi

11	3104	81.9	617	22	AAG78827 Murine SKIP-3. Mu
12	2875	75.8	659	23	ABB04797 LDL receptor bindi
13	2864	75.5	659	23	ABB04793 LDL receptor bindi
14	2846	75.1	659	23	ABB04790 LDL receptor bindi
15	2845	75.0	659	23	ABB04794 LDL receptor bindi
16	2829	74.6	659	23	ABB04789 LDL receptor bindi
17	2824	74.5	659	23	ABB04791 LDL receptor bindi
18	2817	74.3	659	23	ABB04792 LDL receptor bindi
19	2816	74.3	659	23	ABB04795 LDL receptor bindi
20	2814	74.2	659	23	ABB04796 LDL receptor bindi
21	1428	37.7	299	20	AAW88550 Secreted protein e
22	1428	37.7	299	22	ABB50317 Human secreted pro
23	1279.5	33.7	830	23	ABB04608 LDL receptor bindi
24	1116.5	29.4	830	23	ABB04813 LDL receptor bindi
25	1111.5	29.3	830	23	ABB04810 LDL receptor bindi
26	1103.5	29.1	829	23	ABB04817 LDL receptor bindi
27	1097.5	28.9	830	23	ABB04815 LDL receptor bindi
28	1093.5	28.8	830	23	ABB04811 LDL receptor bindi
29	1083.5	28.6	830	23	ABB04809 LDL receptor bindi
30	1075.5	28.4	830	23	ABB04814 LDL receptor bindi
31	1074	28.3	830	23	ABB04816 LDL receptor bindi
32	1060.5	28.0	830	23	ABB04812 LDL receptor bindi
33	1056.5	27.9	598	20	AAW99065 Human U62317 prote
34	751	19.8	155	20	AAW81527 JIP-1 JNK binding
35	537.5	14.2	483	22	AAW57918 Drosophila melanog
36	392	10.3	79	20	AAW81537 JIP-1 polypeptide
37	389	10.3	77	20	AAW81538 JIP-1 polypeptide
38	359	9.5	76	20	AAW81536 JIP-1 polypeptide
39	327	8.6	68	20	AAW81539 JIP-1 polypeptide
40	305	8.0	59	20	AAW81540 JIP-1 polypeptide
41	261	6.9	49	20	AAW81541 JIP-1 polypeptide
42	260	6.9	50	20	AAW81535 JIP-1 SH3 domain
43	254	6.7	66	22	ABB50572 Human secreted pro
44	207	5.5	39	20	AAW81542 JIP-1 polypeptide
45	206	5.4	1239	22	ABG09877 Novel human diagno

#### ALIGNMENTS

#### RESULT 1

AAW80602  
ID AAW80602 standard; Protein; 711 AA.

XX AAW80602;

DT 15-FEB-1999 (first entry)

DE Human transcription factor islet-brain 1 (IB1).

XX IBI, islet-brain 1; transcription factor; human; diabetes;  
dementia; Parkinson's disease; Alzheimer's disease; epilepsy;  
neuroblastoma; glioblastoma; apoptosis; cancer; autoimmune disease;  
systemic lupus erythematosus; myocardial infarction; ischaemia;  
diagnosis; therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 282

FT /note= "encoded by TAC"

XX WO9844106-A1.

XX 08-OCT-1998.

XX 02-APR-1998; 98WO-GB00972.

XX 15-MAY-1997; 97GB-0009920.

XX 03-APR-1997; 97GB-0006731.

XX (KIDD/) KIDDLE S J.

XX (NICO/) NICOD P.

PA (WAEB/) WAEBER G.  
XX  
PI Bonny C, Waeber G;  
XX  
XX WPI: 1998-568278/48.  
DR N-PSDB: AAV62463.  
XX  
XX New isolated transcription factor islet-brain 1 - used to develop  
PT products for treating e.g. diabetes, neurodegenerative disorders,  
PT cancers, autoimmune disease, heart disease or epilepsy  
XX  
XX Claim 2; Fig 1F; 11pp; English.  
PS  
XX  
XX This is the amino acid sequence of human islet-brain 1 (IB1), a  
CC novel transcriptional activator that is involved in the control of  
CC the GLUT2 and insulin genes by interacting with homologous  
CC cis-regulatory elements of the GLUT2 and insulin gene promoters.  
CC The amino acid sequence was deduced from the sequence of isolated  
CC IB1 cDNA (see AAV80602) derived from human insulinoma INS-1 cells.  
CC Rat IB1 polynucleotide (see AAV62462) and polypeptide (see AAV80601)  
CC are also claimed. IB1 polypeptides, nucleic acids, agonists and  
CC antagonists can be used in the treatment or diagnosis of diabetes,  
CC neurodegenerative disorders or cancers, to inhibit cells from  
CC undergoing apoptosis, to treat dementia, Parkinsonism, Alzheimer's  
CC disease, neuronal disabilities such as speech disorders and memory  
CC alteration, autoimmune diseases affecting the CNS such as systemic  
CC lupus erythematosus, diabetes, heart diseases such as myocardial  
CC infarct and ischemia or brain attack, neurological disorders, e.g.  
CC neuroblastoma, glioblastoma or cancers, or to promote apoptosis in  
CC cells, or treat refractory epilepsy. They can also be used for  
CC screening test substances for IB1 biological activity. IB1  
CC producing cells can be used therapeutically to produce IB1 in a  
CC subject.  
XX  
SQ Sequence 711 AA;

Query Match 99.7%; Score 3780; DB 19; Length 711;  
Best Local Similarity 99.7%; Pred No. 4.4e-287;  
Matches 709; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAERESGLGGAASPPAASPLGLHIAAPPNFRTHDISEEFEDLSEITDCGSL 60  
DB 1 MAERESGLGGAASPPAASPLGLHIAAPPNFRTHDISEEFEDLSEITDCGSL 60

QY 61 QCKDTLSLRRPRAGLLSAGGGAGSRLQAEMLQMDLIDATGTPGAEDDEDDDEERAR 120  
DB 61 QCKDTLSLRRPRAGLLSAGGGAGSRLQAEMLQMDLIDATGTPGAEDDEDDDEERAR 120

QY 121 RCGAGPPKAESGQEPASRGQSGQSGDYRPRKPTTLNLFQVPRSQDTLNNN 180  
DB 121 RCGAGPPKAESGQEPASRGQSGQSGQSGDYRPRKPTTLNLFQVPRSQDTLNNN 180

QY 181 SLGKHSWQDRVSRSSPLKGTQTPPHEHICLSELPQSGPAPTDRGTSTDSPCRRS 240  
DB 181 SLGKHSWQDRVSRSSPLKGTQTPPHEHICLSELPQSGPAPTDRGTSTDSPCRRS 240

QY 241 TATQMAPPGGPPAAPGGGRHSRDRHYQADVRLATEEILYTPVQRPDAEPTSAFL 300  
DB 241 TATQMAPPGGPPAAPGGGRHSRDRHYQADVRLATEEILYTPVQRPDAEPTSAFL 300

QY 301 PPTESRMSVSDPDPAAYPSTAGRPHPSISEEEGFDCLSSPERAPPPGGWGRGSLGEPP 360  
DB 301 PPTESRMSVSDPDPAAYPSTAGRPHPSISEEEGFDCLSSPERAPPPGGWGRGSLGEPP 360

QY 361 PPRASLSSDTSALSVDKVTYLVDEHAQLELVSRPCFGYSDSDSATYVDCASYS 420  
DB 361 PPRASLSSDTSALSVDKVTYLVDEHAQLELVSRPCFGYSDSDSATYVDCASYS 420

QY 421 SPYESATGEEVEAPRQPPACLSSESTDPDPVHFSKFLNVFMGRSRSSAESFGFL 480  
DB 421 SPYESATGEEVEAPRQPPACLSSESTDPDPVHFSKFLNVFMGRSRSSAESFGFL 480

QY 481 SCIIINGEEQETHRAIFRFPVRHEDELEVEDDPLLVQLAEDYWEAYNMRTGARGVFP 540

DB 481 SCIIINGEEQETHRAIFRFPVRHEDELEVEDDPLLVQLAEDYWEAYNMRTGARGVFP 540  
QY 541 AYAIAEVTKEPEHMAALAKNSDWQFRVKFLGSGVQVPHKGNVLCAMOKIATTRLT 600  
DB 541 AYAIAEVTKEPEHMAALAKNSDWQFRVKFLGSGVQVPHKGNVLCAMOKIATTRLT 600  
QY 601 VHENPPSSCVLETISVRGVKIGVKADDSDQAKGNKCSHFFQFNISFCGYHPKNNKYFGFI 660  
DB 601 VHENPPSSCVLETISVRGVKIGVKADDSDQAKGNKCSHFFQFNISFCGYHPKNNKYFGFI 660  
QY 661 TKHPADHRFACHVFVSESDTKALAESVGRFAFOFYKQFVEYTCPTEDIYLE 711  
DB 661 TKHPADHRFACHVFVSESDTKALAESVGRFAFOFYKQFVEYTCPTEDIYLE 711

RESULT 2  
AAAG78707  
ID AAG78707 standard; Protein; 708 AA.  
XX  
AC AAG78707;  
XX  
DT 11-DEC-2001 (first entry)  
XX  
DE Rat SAPK interacting protein #2.  
XX  
KW Rat; SKIP; SAPK interacting protein; stress-activated kinase;  
KW cancer; inflammation; apoptosis; cell death; apoptosis.  
XX  
OS Rattus norvegicus.  
XX  
FH Key Location/Qualifiers  
FT Domain 126..283  
FT /label= JNK\_binding\_domain  
FT Domain 366..370  
FT /label= proline\_rich\_domain  
FT Domain 484..546  
FT /label= SH3\_domain  
FT Domain 529..594  
FT /label= helix-loop-helix\_motif  
FT Domain 563..697  
FT /label= phosphotyrosine\_interaction\_domain  
XX  
PN KR2001029353-A.  
XX  
XX 06-APR-2001.  
XX  
XX 30-SEP-1999; 99KR-0042119.  
XX  
XX 30-SEP-1999; 99KR-0042119.  
XX  
PA (CHOI/) CHOI E J.  
XX  
PI Choi EJ, Han PR, Kim IJ, Park BY, Park JH;  
XX  
XX WPI: 2001-569558/64.  
XX N-PSDB: AAI66505.  
XX  
XX Rat's skip (sapk interacting protein) selectively inhibiting  
PT stress-activated protein kinase and gene thereof -  
XX  
XX Claim 2; Page 20-24; 34pp; Korean.  
XX  
XX The present invention provides the protein and coding sequences of  
CC different versions of the rat stress-activated kinase (SAPK) interacting  
CC protein (SKIP). This protein is capable of modulating cell death, and  
CC can be used in the treatment of cancer, inflammation, apoptosis and  
CC similar diseases. The present sequence is one version of the protein  
CC of the invention.  
XX  
SQ Sequence 708 AA;

Query Match 93.5%; Score 3544.5; DB 22; Length 708;

Best Local Similarity 94.0%; Pred. No. 1.2e-268;  
Matches 669; Conservative 11; Mismatches 27; Indels 5; Gaps 4;

QY 1 MAERESGLGGGAASPPAASPFGLGLHIASPPNFRTHDISLEEFEDDLSEITDECGISL 60  
DB 1 MAERES-GLSGGAASPPAASPFGLGLHIASPPNFRTHDISLEEFEDDLSEITDECGISL 59  
QY 61 OCKDTLSLRPRAGLLSAG-GGGAGSRLOAEMQLMDLIDATGTPGAEDDEDDDERAA 119  
DB 60 OCKDTLSLRPRAGLLSAGSGSAGSRLOAEMQLMDLIDAASTPGAEDDEDD-ELAA 118  
QY 120 RPPGAGPPKAESGOEPASRGQSGQSGQSGDYTKRPTTLNLPQVPRSQDTLNN 179  
DB 119 RPPGAGPPKAESGOEPASRGQ--QGQGGTGGDYTKRPTTLNLPQVPRSQDTLNN 176  
QY 180 NSLGGKHSWQDRVSRSSPLKTGEQTPPHEHICLSDELPPQSGPAPTDRGTSTDSPCR 239  
DB 177 NSLGGKHSWQDRVSRSSPLKTGEQTPPHEHICLSDELPPQSGPAPTDRGTSTDSPCR 236  
QY 240 STATOMAPPGPPAAPGCGHSHRDRIHQADVLEATEEIIYLTVPORPDAAETSAF 299  
DB 237 TAATOMAPPGPPAATAPGCGHSHRDRIHQADVLEATEEIIYLTVPORPDAAETSTF 296  
QY 300 LPTESRMSVSSDDPPAAYSTAGRPHPSISEEEGFDCLSSPERAPPGGGWRGSLGEP 359  
DB 297 LPTESRMSVSSDDPPAAYSVTAGRPHPSISEEEGFDCLSSPERAPPGGGWRGSLGEP 356  
QY 360 PPPRASLSSDTSALSVDYKTYLVVDEHAQLELVSLRCPFGDYSDSDSATVYDNCASV 419  
DB 357 PPPRASLSSDTSALSVDYKTYLVVDEHAQLELVSLRCPFGDYSDSDSATVYDNCASA 416  
QY 420 SSPYSAIGEYEEAPRPPACLSDESDTPDEPDVHFSSKFLNFMGSRSSSAESFGL 479  
DB 417 SSPYSAIGEYEEAPRPPACLSDESDTPDEPDVHFSSKFLNFMGSRSSSAESFGL 476  
QY 480 FSCINGEEOQTHRAIFRVPVPRHEDELEVDPLLVQLAEADYWEAYNMRTGARGVF 539  
DB 477 FSCINGEEOQTHRAIFRVPVPRHEDELEVDPLLVQLAEADYWEAYNMRTGARGVF 536  
QY 540 PAYVAIEVTRKEPMAALAKNSDWQFRVFKLGSVQVYPYHKGNDVLCAMQKIATTRL 599  
DB 537 PAYVAIEVTRKEPMAALAKNSDWIDQFRVFKLGSVQVYPYHKGNDVLCAMQKIATTRL 596  
QY 600 TVHENPSSCVLEISVGVKIGVAKDDSOEAKGKCSHFFOLKNISCGYHPKNNKYEGF 659  
DB 597 TVHENPSSCVLEISVGVKIGVAKDEAQAEGKNCSEHFFOLKNISCGYHPKNNKYEGF 656  
QY 660 ITKHPADHRFACHVFSSEDTKALAESVGRFAQFQYKOFVEYTCPTEDIYLE 711  
DB 657 ITKHPADHRFACHVFSSEDTKALAESVGRFAQFQYKOFVEYTCPTEDIYLE 708

## RESULT 3

AAG78708

ID AAG78708 standard; Protein; 715 AA.

XX AAG78708;

AC AAG78708;

XX AAG78708;

DT 11-DEC-2001 (first entry)

XX Rat SAPK interacting protein #3.

DE DE

KW Rat; SKIP; SAPK interacting protein; stress-activated kinase;  
KW cancer; inflammation; apoptosis; cell death; apoptosis.

XX Rattus norvegicus.

OS

XX Location/Qualifiers

FH Key 133..290

FT Domain /label= JNK\_binding\_domain

FT Domain 373..377

FT Domain /label= proline\_rich\_domain

FT Domain 491..553

FT Domain /label= SH3\_domain  
FT 536..601  
FT /label= helix-loop-helix\_motif  
FT 570..703  
FT /label= phosphotyrosine\_interaction\_domain  
XX  
PN KR2001029353-A.

XX 06-APR-2001.

XX 30-SEP-1999; 99KR-0042119.

XX 30-SEP-1999; 99KR-0042119.

XX (CHOI/) CHOI E J.

XX Choi EJ, Han PR, Kim IJ, Park BY, park JH;

XX WPI: 2001-569558/64.

XX N-PSDB; AAI66506.

XX Rat's skip (sapk interacting protein) selectively inhibiting

XX stress-activated protein kinase and gene thereof -

XX Claim 2; Page 27-30; 34pp; Korean.

XX The present invention provides the protein and coding sequences of

XX different versions of the rat stress-activated kinase (SAPK) interacting

XX protein (SKIP). This protein is capable of modulating cell death, and

XX can be used in the treatment of cancer, inflammation, apoptosis and

XX similar diseases. The present sequence is one version of the protein

XX of the invention.

XX SQ Sequence 715 AA;

Query Match 93.5%; Score 3544.5; DB 22; Length 715;

Best Local Similarity 94.0%; Pred. No. 1.2e-268;

Matches 669; Conservative 11; Mismatches 27; Indels 5; Gaps 4;

QY 1 MAERESGLGGGAASPPAASPFGLGLHIASPPNFRTHDISLEEFEDDLSEITDECGISL 60

DB 8 MAERES-GLSGGAASPPAASPFGLGLHIASPPNFRTHDISLEEFEDDLSEITDECGISL 66

QY 61 OCKDTLSLRPRAGLLSAG-GGGAGSRLOAEMQLMDLIDATGTPGAEDDEDDDERAA 119

DB 67 OCKDTLSLRPRAGLLSAGSGSAGSRLOAEMQLMDLIDAASTPGAEDDEDD-ELAA 125

QY 120 RPPGAGPPKAESGOEPASRGQSGQSGQSGDYTKRPTTLNLPQVPRSQDTLNN 179

DB 126 RPPGAGPPKAESGOEPASRGQ--QGQGGTGGDYTKRPTTLNLPQVPRSQDTLNN 183

QY 180 NSLGGKHSWQDRVSRSSPLKTGEQTPPHEHICLSDELPPQSGPAPTDRGTSTDSPCR 239

DB 184 NSLGGKHSWQDRVSRSSPLKTGEQTPPHEHICLSDELPPQSGPAPTDRGTSTDSPCR 243

QY 240 STATOMAPPGPPAAPGCGHSHRDRIHQADVLEATEEIIYLTVPORPDAAETSAF 299

DB 244 TAATOMAPPGPPAATAPGCGHSHRDRIHQADVLEATEEIIYLTVPORPDAAETSTF 303

QY 300 LPTESRMSVSSDDPPAAYSTAGRPHPSISEEEGFDCLSSPERAPPGGGWRGSLGEP 359

DB 304 LPTESRMSVSSDDPPAAYSVTAGRPHPSISEEEGFDCLSSPERAPPGGGWRGSLGEP 363

QY 360 PPPRASLSSDTSALSVDYKTYLVVDEHAQLELVSLRCPFGDYSDSDSATVYDNCASV 419

DB 364 PPPRASLSSDTSALSVDYKTYLVVDEHAQLELVSLRCPFGDYSDSDSATVYDNCASA 423

QY 420 SSPYSAIGEYEEAPRPPACLSDESDTPDEPDVHFSSKFLNFMGSRSSSAESFGL 479

DB 424 SSPYSAIGEYEEAPRPPACLSDESDTPDEPDVHFSSKFLNFMGSRSSSAESFGL 483

QY 480 FSCINGEEOQTHRAIFRVPVPRHEDELEVDPLLVQLAEADYWEAYNMRTGARGVF 539

Db 484 FSCVINGEEHOTHRAIFRVPVPRHEDELELEVDOPLLVELQAEDYWEAYNNMRTGARGVF 543  
 QY 540 PAYIAIEVTKPEPHMAALAKNSDWDOFRVFLGSGVQVPHKGNVLCAMOKIATTRL 599  
 Db 544 PAYIAIEVTKPEPHMAALAKNSDWDOFRVFLGSGVQVPHKGNVLCAMOKIATTRL 603  
 QY 600 TVHFNPPSSCVLEISVRGVKIGVAKDDSOEAKGNKCSHFFOLKNISPCGVHPKNNKYFGF 659  
 Db 604 TVHFNPPSSCVLEISVRGVKIGVAKDEAQEAKGNKCSHFFOLKNISPCGVHPKNNKYFGF 663  
 QY 660 ITKHPADHRFACHVFVSEDSKALAESVGRFAQFOFYKQFVEYTCPTEDIYLE 711  
 Db 664 ITKHPADHRFACHVFVSEDSKALAESVGRFAQFOFYKQFVEYTCPTEDIYLE 715

RESULT 4  
 AAG78825 standard; Protein; 707 AA.  
 XX  
 AC AAG78825;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Murine SKIP-1b.  
 XX  
 KW Murine; SKIP-1b; SAPK-interacting protein; phosphorylase; JNK; SAPK;  
 KW c-Jun N-terminal kinase; Stress-Activated Protein Kinase;  
 KW cell death related disease.  
 XX  
 OS Mus musculus.  
 XX  
 PN KR2001029352-A.  
 XX  
 PD 06-APR-2001.  
 XX  
 PF 30-SEP-1999; 99KR-0042118.  
 XX  
 PR 30-SEP-1999; 99KR-0042118.  
 XX  
 PA (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.  
 XX  
 PI Choi IY, Choi UJ, Han PR, Lee GU, Lee JG, Lee SH;  
 XX  
 DR WPI; 2001-563617/63.  
 XX  
 DR N-PSDB; AAI65090.  
 XX  
 PT Mouse skip (SAPK-interacting protein) controlling activation of human  
 PT cell phosphorylase, jnk, and its gene, useful for treating cell death  
 PT related diseases -  
 XX  
 PS Claim 2; Page 18-21; 37pp; Korean.  
 XX  
 CC The present sequence is the protein sequence for murine skip-1b  
 CC (SAPK-interacting protein), which selectively inhibits or controls the  
 CC activation of a phosphorylase, JNK/SAPK (c-Jun N-terminal kinase/  
 CC Stress-Activated Protein Kinase) activated by various kinds of stress.  
 CC SKIP and its gene are useful in treating cell death related diseases.  
 XX  
 SQ Sequence 707 AA;

Query Match 93.1%; Score 3529; DB 22; Length 707;  
 Best Local Similarity 93.3%; Pred. No. 1.9e-267;  
 Matches 665; Conservative 13; Mismatches 27; Indels 8; Gaps 4;

QY 1 MARESGGLGGGAAPPAAAPPLGLHIASPPNFRLLTHDISEEFEDLSLETDECGISL 60  
 Db 1 MARESGGLGGGAAPPAAAPPLGLHIASPPNFRLLTHDISEEFEDLSLETDECGISL 59  
 QY 61 QCKDITLSLRPRAGLLSAG-GGAGSRQLQAEMLQMDLIDATGTPGAEDDEDDDEERA 119  
 Db 60 QCKDITLSLRPRAGLLSAGSGSRLQAEMLQMDLIDAGTTPGAEDDEDDDEELAA 119  
 QY 120 RPPGAGPPKAESQBPASRGQSQGSGDGYRPRPTTLNLFQVPRSQDTLNN 179

Db 120 QRCVGPAPKASNQDPAPRSQOG-----ATGSGDTYRPRPTTLNLFQVPRSQDTLNN 174  
 QY 180 NSLGKKHNSQDVRVSRSSSPLTKGTQTPPHEHICLSBELPQSGPAPATTDRGTSTDSPCR 239  
 Db 175 NSLGKKHNSQDVRVSRSSSPLTKGTQTPPHEHICLSBELPQSGPVPQDRGTSTDSPCR 234  
 QY 240 STATOMAPGGPAPPGGGRGSHRDRIHQADVRLATEEIEVLTVPQRPDPAEPTSF 299  
 Db 235 SAATOMAPSGPATAPGGGRGSHRDRIHQADVRLATEEIEVLTVPQRPDPAEPTSF 294  
 QY 300 LPPTESRMVSSDDPAPAYPTAGRPHPISIEEEGFDCLSSPERAEPGGGRCGLGP 359  
 Db 295 MPPTESRMVSSDDPAPAYSVTAGRPHPISIEEEGFDCLSSPERAEPGGGRCGLGP 354  
 QY 360 PPPRASLSSTALSYSYVKTLLVVDHAQLELVSLRPGFCGYSDESATSATYDNCASV 419  
 Db 355 PPPRASLSSTALSYSYVKTLLVVDHAQLELVSLRPGFCGYSDESATSATYDNCASA 414  
 QY 420 SSPYESAIGEYEEAPRPPQACLSDESTDPDPVHFSSKFLNVFMVSGRSRSSAESFGL 479  
 Db 415 SSPYESAIGEYEEAPRPPQACLSDESTDPDPVHFSSKFLNVFMVSGRSRSSAESFGL 474  
 QY 480 FSCINGEEOETHRAIFRVPVPRHEDELELEVDOPLLVELQAEDYWEAYNNMRTGARGVF 539  
 Db 475 FSCVINGEEOETHRAIFRVPVPRHEDELELEVDOPLLVELQAEDYWEAYNNMRTGARGVF 534  
 QY 540 PAYIAIEVTKPEPHMAALAKNSDWDOFRVFLGSGVQVPHKGNVLCAMOKIATTRL-R 598  
 Db 535 PAYIAIEVTKPEPHMAALAKNSDWDOFRVFLGSGVQVPHKGNVLCAMOKIATTRLR 594  
 QY 599 LTVHFNPPSSCVLEISVRGVKIGVAKDDSOEAKGNKCSHFFOLKNISPCGVHPKNNKYFG 658  
 Db 595 LTVHFNPPSSCVLEISVRGVKIGVAKDDSOEAKGNKCSHFFOLKNISPCGVHPKNNKYFG 654  
 QY 659 FITKHPADHRFACHVFVSEDSKALAESVGRFAQFOFYKQFVEYTCPTEDIYLE 711  
 Db 655 FITKHPADHRFACHVFVSEDSKALAESVGRFAQFOFYKQFVEYTCPTEDIYLE 707

RESULT 5  
 AAW80601 standard; Protein; 714 AA.  
 XX  
 AC AAW80601;  
 XX  
 DT 15-FEB-1999 (first entry)  
 XX  
 DE Rat transcription factor islet-brain 1 (IB1).  
 XX  
 KW IB1; islet-brain 1; transcription factor; rat; diabetes;  
 KW dementia; Parkinson's disease; Alzheimer's disease; epilepsy;  
 KW neuroblastoma; glioblastoma; apoptosis; cancer; autoimmune disease;  
 KW systemic lupus erythematosus; myocardial infarction; ischaemia;  
 KW diagnosis; therapy.  
 XX  
 OS Rattus sp.  
 XX  
 FH Key  
 FT Region 31..61  
 FT /note= "alpha-helical structure"  
 FT Region 114..125  
 FT /note= "alpha-helical structure"  
 FT Region 292..366  
 FT /note= "proline-rich region"  
 FT Peptide 163..190  
 FT /note= "putative nuclear localisation signal"  
 FT Peptide 242..270  
 FT /note= "putative nuclear localisation signal"  
 PN W09844106-A1.  
 XX  
 PD 08-OCT-1998.

XX PF 02-APR-1998; 98WO-GB000972.  
 XX PR 15-MAY-1997; 97GB-0009920.  
 XX PR 03-APR-1997; 97GB-0006731.  
 XX PA (KIDD/) KIDDLE S J.  
 XX PA (NICOD/) NICOD P.  
 XX PA (WAEB/) WAEBER G.  
 XX PI Bonny C, Waerber G;  
 XX DR WPI: 1998-568278/48.  
 XX DR N-PSDB; AAV62462.  
 XX  
 CC New isolated transcription factor islet-brain 1 - used to develop  
 PT products for treating e.g. diabetes, neurodegenerative disorders,  
 PT cancers, autoimmune disease or epilepsy  
 XX  
 PS Claim 1; Fig 1A; 111pp; English.  
 XX  
 CC This is the amino acid sequence of rat islet-brain 1 (IB1), a novel  
 CC transcriptional activator that is involved in the control of the  
 CC GLUT2 and insulin genes by interacting with homologous  
 CC cis-regulatory elements of the GLUT2 and insulin gene promoters.  
 CC The amino acid sequence was deduced from the sequence of isolated  
 CC IB1 cDNA (see AAV80601) derived from rat insulinoma INS-1 cells.  
 CC Human IB1 polynucleotide (see AAV62463) and polypeptide (see AAV80602)  
 CC are also claimed. IB1 polypeptides, nucleic acids, agonists and  
 CC antagonists can be used in the treatment or diagnosis of diabetes,  
 CC neurodegenerative disorders or cancers, to inhibit cells from  
 CC undergoing apoptosis, to treat dementia, Parkinsonism, Alzheimer's  
 CC disease, neuronal disabilities such as speech disorders and memory  
 CC alteration, autoimmune diseases affecting the CNS such as systemic  
 CC lupus erythematosus, diabetes, heart diseases such as myocardial  
 CC infarct and ischemia or brain attack, neurological disorders, e.g.  
 CC neuroblastoma, glioblastoma or cancers, or to promote apoptosis in  
 CC cells, or treat refractory epilepsy. They can also be used for  
 CC screening test substances for IB1 biological activity. IB1  
 CC producing cells can be used therapeutically to produce IB1 in a  
 CC subject.  
 XX  
 SQ Sequence 714 AA;  
 Query Match 92.4%; Score 3503; DB 19; Length 714;  
 Best Local Similarity 93.3%; Pred. No. 2.1e-265;  
 Matches 664; Conservative 11; Mismatches 31; Indels 6; Gaps 5;  
 QY 1 MAERESGLGGGAAPPAASPLGLHTASPPNFRTHDLSLEEFDEDLSEITDECGISL 60  
 DB 8 MAERES-GLSGGAAPPAASPLGLHTASPPNFRTHDLSLEEFDEDLSEITDECGISL 66  
 QY 61 QCKDTLSLRPPRAGLSAG-GGGAGSLQAEMQLMDLIDATGTPGAEDDEEDDERAA 119  
 DB 67 QCKDTLSLRPPRAGLSAGSGSGLQAEMQLMDLIDASDTPGAEDDEEDD-ELAA 125  
 QY 120 RRGAGPKAESGQEPASRGQSGQSGQSGDTPYRKRPTTLNLFQVPRSQDTLNN 179  
 DB 126 QRPVGPSKAESGQEPASRSQ--QGQPGTGCQDTPYRKRPTTLNLFQVPRSQDTLNN 183  
 QY 180 NSLGKHKSHQDVRSSSPLTKGTPTPHEHICLSEELPQSGAPATTDRTGSTDSCRR 239  
 DB 184 NSLGKHKSHQDVRSSSPLTKGTPTPHEHICLSEELPQSGAPATTDRTGSTDSCRR 243  
 QY 240 STATQMAPPGGPAAPPGGRGSHDRTHYQADYRLATEEIIYLPVQRPDAEPTSAF 299  
 DB 244 TATQMAPPGGPAAPPGGRGSHDR-SISADVRLATEEIIYLPVQRPDAEPTSTF 302  
 QY 300 LPPTESRMVSSDDPPAAYPTAGRPSPSISEEEGFDCLSPERAEPPGGMRGSLGEP 359  
 DB 303 LPPTESRMVSSDDPPAAYSVTAGRPSPSISEEEDGFDCLSPERAEPPGGMRGSLGEP 362  
 QY 360 PPPRASLSSDTSALSYDSVKTYLVVDEHAQLVLSLRPCFGDYSDESATVYDNCASV 419

DB 363 PPPRASLSSDTSALSYDSVKTYLVVDEHAQLVLSLRPCFGDYSDESATVYDNCASA 422  
 QY 420 SSSPYESAIGEYEEAPRQPPACLSDESTDPEDVHFSSKKFLNVFMSGRSSSAESFGL 479  
 DB 423 SSSPYESAIGEYEEAPRQPPACLSDESTDPEDVHFSSKKFLNVFMSGRSSSAESFGL 482  
 QY 480 FSCIINGEEQETHRAIFRVPRHDELELEVDPLLLVELQAEDYHAYNNMRTGARGVF 539  
 DB 483 FSCVINGEEHEQTHRAIFRVPRHDELELEVDPLLLVELQAEDYHAYNNMRTGARGVF 542  
 QY 540 PAYAIEVTKEPEHMAALAKNSDWDOFRVKFLGSGVQVPYHKGNDVLCAMQKIATTRL 599  
 DB 543 PAYAIEVTKEPEHMAALAKNSDWDOFRVKFLGSGVQVPYHKGNDVLCAMQKIATTRL 602  
 QY 600 TVHNPSSSCVLEISVRGVKIGVADDSQAKNCKSHFFQLKNISFCGYHPKNNKYFGF 659  
 DB 603 TVHNPSSSCVLEISVRGVKIGVADDSQAKNCKSHFFQLKNISFCGYHPKNNKYFGF 662  
 QY 660 ITKHPADHRFACHVFVSEDTKALAESVGRFAQFOQYKQFVEYTCPTEDIYLE 711  
 DB 663 ITKHPADHRFACHVFVSEDTKALAESVGRFAQFOQYKQFVEYTCPTEDIYLE 714  
 RESULT 6  
 AAG78706  
 ID AAG78706 standard; Protein; 699 AA.  
 XX AC AAG78706;  
 XX DT 11-DEC-2001 (first entry)  
 XX DE Rat SAPK interacting protein #1.  
 XX KW Rat; SKIP; SAPK interacting protein; stress-activated kinase;  
 XX KW cancer; inflammation; apoptosis; cell death; apoptosis.  
 XX OS Rattus norvegicus.  
 XX FH Key Location/Qualifiers  
 XX FT Domain 117..274  
 XX FT Domain /label= JNK\_binding\_domain  
 XX FT Domain 357..361  
 XX FT Domain /label= proline\_rich\_domain  
 XX FT Domain 475..537  
 XX FT Domain /label= SH3\_domain  
 XX FT Domain 520..585  
 XX FT Domain /label= helix-loop-helix\_motif  
 XX FT Domain 554..688  
 XX FT /label= phosphotyrosine\_interaction\_domain  
 XX PN KR2001029353-A.  
 XX PD 06-APR-2001.  
 XX PF 30-SEP-1999; 99KR-0042119.  
 XX PR 30-SEP-1999; 99KR-0042119.  
 XX PA (CHOI/) CHOI E J.  
 XX PI Choi EJ, Han PR, Kim IJ, Park BY, Park JH;  
 XX DR WPI: 2001-569558/64.  
 XX DR N-PSDB; AAI66504.  
 XX PT Rat's skip (sapk interacting protein) selectively inhibiting  
 XX PT stress-activated protein kinase and gene thereof -  
 XX PS Claim 2; Page 14-17; 34pp; Korean.  
 XX CC The present invention provides the protein and coding sequences of  
 XX CC different versions of the rat stress-activated kinase (SAPK) interacting

CC protein (SKIP). This protein is capable of modulating cell death, and  
CC can be used in the treatment of cancer, inflammation, apoplexy and  
CC similar diseases. The present sequence is one version of the protein  
CC of the invention.

XX	SQ	Sequence	699 AA;
XX	Query Match	89.4%; Score 3389; DB 22; Length 699;	
XX	Best Local Similarity	94.0%; Pred. No. 1.7e-256;	
XX	Matches	637; Conservative 11; Mismatches 26; Indels 4; Gaps 3;	
Qy	35	LTHDISLEEFDEDLSETDECGISLOCKDTLSRPPRAGLLSAG-GGGAGSRLQAEMLQ	93
Db	25	LTHDISLEEFDEDLSETDECGISLOCKDTLSRPPRAGLLSAGSGSAGSRLQAEMLQ	84
Qy	94	MDLIDATGDTGAEDDEDDDEERAARRPGAGPPKAESGQEPASRGOGSQSGPGSG	153
Db	85	MDLIDAAADTPGAEDDEDD-ELAAQRPGVGPKAESGQEPASRGOG--QGQPGTGS	141
Qy	154	DTYRKRPTTLNLFPOVPRSDTLNNSLGKHSQWDRVSRSSSPLKTGOTPPHEHICL	213
Db	142	DTYRKRPTTLNLFPOVPRSDTLNNSLGKHSQWDRVSRSSSPLKTGOTPPHEHICL	201
Qy	214	SEELPPQSGPAPTTDRGTSTDSPCRSTATQMAPGGPPAAPPGGRGHSHRDRIHQADV	273
Db	202	SDELPPQSGPVPQTDRGTSTDSPCRSTAAQTQMAPSGPPATAPGGRGHSHRDRIHQADV	261
Qy	274	RLEATEEILYTPVQRPPDAEPTSAFLPPTESRMSVSDPDPAAPYPTAGRPHPSISEE	333
Db	262	RLEATEEILYTPVQRPPDAEPTSTFLPPTESRMSVSDPDPAAPYPTAGRPHPSISEE	321
Qy	334	EGFDCLSSEPERAEPGGGWRGSLGEPPPPRASTSSDTLSALSYDSVKYTLVVDHAQLEL	393
Db	322	EGFDCLSSEPERAEPGGGWRGSLGEPPPPRASTSSDTLSALSYDSVKYTLVVDHAQLEL	381
Qy	394	VSLRPFQDYSDESATVYDNCASVSPYSAIGEYEEAPRPPACLSDESTPDEPD	453
Db	382	VSLRPFQDYSDESATVYDNCASVSPYSAIGEYEEAPRPPACLSDESTPDEPD	441
Qy	454	VHFSKFLNVFMGSRSSSAESFGLFSCVINGEEOQTHRAIFRVPRHEDELEVEDD	513
Db	442	VHFSKFLNVFMGSRSSSAESFGLFSCVINGEEOQTHRAIFRVPRHEDELEVEDD	501
Qy	514	PLLVELQAEADWYEAYNMRTGARGVFPAYIAEVTKEPHEMAALAKNSDWQDQFRVKFLG	573
Db	502	PLLVELQAEADWYEAYNMRTGARGVFPAYIAEVTKEPHEMAALAKNSDWIDQFRVKFLG	561
Qy	574	SVQVPYHKGNDVLCAMQKIATTRLTVHFNPPSSCVLEISVRGVKIGVKADDSQEAQGN	633
Db	562	SVQVPYHKGNDVLCAMQKIATTRLTVHFNPPSSCVLEISVRGVKIGVKADDEQAQGN	621
Qy	634	KCSHFPQKNISFCGYHPKNNKYFGITKHPADHRFACHVFVSEDSTKALAESVGRAFOQ	693
Db	622	KCSHFPQKNISFCGYHPKNNKYFGITKHPADHRFACHVFVSEDSTKALAESVGRAFOQ	681
Qy	694	FYKQFVETCTPTEDIYLE 711	
Db	682	FYKQFVETCTPTEDIYLE 699	

RESULT 7  
AAG78826  
ID AAG78826 standard; Protein; 698 AA.  
XX AAG78826;  
AC AAG78826;  
XX  
DT 18-DEC-2001 (first entry)  
XX Murine SKIP-2a.  
XX Murine; SKIP-2a; SAPK-interacting protein; phosphotyrase; JNK; SAPK;  
KW c-Jun N-terminal kinase; Stress-Activated Protein Kinase;  
KW cell death related disease.

XX	Mus musculus.		
XX	KR2001029352-A.		
XX	06-APR-2001.		
XX	30-SEP-1999; 99KR-00421118.		
XX	30-SEP-1999; 99KR-00421118.		
XX	(KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.		
XX	Choi IV, Choi UU, Han PR, Lee GU, Lee JG, Lee SH;		
DR	WPI; 2001-563617/63.		
DR	N-PSDB; AA165091.		
XX	Mouse skip (SAPK-interacting protein) controlling activation of human		
PT	cell phosphotyrase, jnk, and its gene, useful for treating cell death		
PT	related diseases -		
XX	Claim 2; Page 24-27; 37pp; Korean.		
XX	The present sequence is the protein sequence for murine SKIP-2a		
CC	(SAPK-interacting protein), which selectively inhibits or controls the		
CC	activation of a phosphotyrase, JNK/SAPK (c-Jun N-terminal kinase/		
CC	Stress-Activated Protein Kinase), activated by various kinds of stress.		
CC	SKIP and its gene are useful in treating cell death related diseases.		
XX	Sequence 698 AA;		
SQ	Query Match 88.8%; Score 3367.5; DB 22; Length 698;		
	Best Local Similarity 93.1%; Pred. No. 8.2e-255;		
	Matches 632; Conservative 13; Mismatches 27; Indels 7; Gaps 3;		

Qy	35	LTHDISLEEFDEDLSETDECGISLOCKDTLSRPPRAGLLSAG-GGGAGSRLQAEMLQ	93
Db	25	LTHDISLEEFDEDLSETDECGISLOCKDTLSRPPRAGLLSAGSGSAGSRLQAEMLQ	84
Qy	94	MDLIDATGDTGAEDDEDDDEERAARRPGAGPPKAESGQEPASRGOGSQSGPGSG	153
Db	85	MDLIDAAAGTGAEDDEDDDEELAAQRPGVGPKAESQMDPAPRSQGG-----ATGSG	139
Qy	154	DTYRKRPTTLNLFPOVPRSDTLNNSLGKHSQWDRVSRSSSPLKTGOTPPHEHICL	213
Db	140	DTYRKRPTTLNLFPOVPRSDTLNNSLGKHSQWDRVSRSSSPLKTGOTPPHEHICL	199
Qy	214	SEELPPQSGPAPTTDRGTSTDSPCRSTATQMAPGGPPAAPPGGRGHSHRDRIHQADV	273
Db	200	SDELPPQSGPVPQTDRGTSTDSPCRSTATQMAPSGPPATAPGGRGHSHRDRIHQADV	259
Qy	274	RLEATEEILYTPVQRPPDAEPTSAFLPPTESRMSVSDPDPAAPYPTAGRPHPSISEE	333
Db	260	RLEATEEILYTPVQRPPDAEPTSTFLPPTESRMSVSDPDPAAPYPTAGRPHPSISEE	319
Qy	334	EGFDCLSSEPERAEPGGGWRGSLGEPPPPRASTSSDTLSALSYDSVKYTLVVDHAQLEL	393
Db	320	EGFDCLSSEPERAEPGGGWRGSLGEPPPPRASTSSDTLSALSYDSVKYTLVVDHAQLEL	379
Qy	394	VSLRPFQDYSDESATVYDNCASVSPYSAIGEYEEAPRPPACLSDESTPDEPD	453
Db	380	VSLRPFQDYSDESATVYDNCASVSPYSAIGEYEEAPRPPACLSDESTPDEPD	439
Qy	454	VHFSKFLNVFMGSRSSSAESFGLFSCVINGEEOQTHRAIFRVPRHEDELEVEDD	513
Db	440	VHFSKFLNVFMGSRSSSAESFGLFSCVINGEEOQTHRAIFRVPRHEDELEVEDD	499
Qy	514	PLLVELQAEADWYEAYNMRTGARGVFPAYIAEVTKEPHEMAALAKNSDWQDQFRVKFLG	573
Db	500	PLLVELQAEADWYEAYNMRTGARGVFPAYIAEVTKEPHEMAALAKNSDWIDQFRVKFLG	559
Qy	574	SVQVPYHKGNDVLCAMQKIATTRLTVHFNPPSSCVLEISVRGVKIGVKADDSQEAQGN	632

```

Db 560 SVQVPYHKGNDVLCAMQKTIATTPRLTVHNPSSCVLEISVRGKIGVKADDALEAKG 619
QY 633 NKCSHFQLNISFCGYPHPKNNKYFGITKHPADHRFACHVFVEDSTKALAESVGRATQ 692
Db 620 NKCSHFQLNISFCGYPHPKNNKYFGITKHPADHRFACHVFVEDSTKALAESVGRATQ 679
QY 693 QFYKQFVEYTCPTEDIYLE 711
Db 680 QFYKQFVEYTCPTEDIYLE 698

RESULT 8
AAW81525
ID AAW81525 standard; Protein; 660 AA.
XX
AC AAW81525;
XX
DT 01-MAR-1999 (first entry)
XX
DE Murine JNK-interacting protein 1 (JIP-1).
XX
KW JIP-1; JNK-interacting protein; c-Jun NH2-terminal kinase;
KW signal transduction; inhibitor; mouse; neurodegenerative disease;
KW Parkinson's disease; Alzheimer's disease; blood clot; stroke;
KW malignancy; cancer; leukaemia; autoimmune disease; inflammation;
KW apoptosis; therapy; diagnosis.
XX
OS Mus sp.
XX
FH Key
FH Domain 127..281
FT /label= JBD
FT /note= "claimed JNK binding domain"
FT Domain 148..174
FT /note= "claimed JBD core"
FT Domain 491..540
FT /label= SH3
XX
PN W09849188-A1.
XX
PD 05-NOV-1998.
XX
PF 28-APR-1998; 98WO-US08513.
XX
PR 28-APR-1997; 97US-0819177.
XX
PA (UYMA-) UNIV MASSACHUSETTS.
XX
PI Davis RJ, Dickens M;
XX
PI WPI: 1999-024042/02.
XX
DR N-PSDB; AAV69289.
XX
XX c-Jun NH2-terminal kinase (JNK)-interacting protein 1 - used to
PT treat neurodegenerative disease, blood clot, leukaemia, autoimmune
PT disease, and inflammation
XX
PS Claim 5; Fig 1B; 95pp; English.
XX
CC This is the amino acid sequence of murine JNK-interacting protein 1
CC (JIP-1), a novel cytoplasmic anchor protein that specifically binds
CC to and inhibits the biological effects of JNK (c-Jun NH2-terminal
CC kinase), including the initiation of apoptosis and oncogenic
CC transformation. The sequence is predicted from a full-length cDNA
CC clone (see AAV69289) isolated from a mouse brain cDNA library. The
CC invention provides JIP-1 nucleic acids and polypeptides (see also
CC AAW81535-45), expression vectors and host cells. The JIP-1
CC polypeptides and nucleic acids (including antisense and ribozymes)
CC can be used in the manufacture of a medicament for treating a
CC pathological condition associated with abnormal expression or
CC activity of JNK, such as a neurodegenerative disease (selected from
CC Parkinson's disease and Alzheimer's disease), a blood clot, stroke,

```

```

CC malignancy, leukaemia, an autoimmune disease or inflammation (all
CC claimed).
XX
SQ Sequence 660 AA;
Query Match 86.0%; Score 3262.5; DB 20; Length 660;
Best Local Similarity 87.1%; Pred. No. 1.3e-246;
Matches 620; Conservative 13; Mismatches 26; Indels 53; Gaps 4;
QY 1 MAERESGLGGCAASPAAAPFLGLHIASPPNFRLLTHDISLEEFEDDLSEITDECGISL 60
Db 1 MAERES-GLGGCAASPAAAPFLGLHIASPPNFRLLTHDISLEEFEDDLSEITDECGISL 59
QY 61 QCKDTLSLRPPRAGLISAG-GGGAGSRLQAEMLQMDLIDATGDTPCAEDDEDDEERAA 119
Db 60 QCKDTLSLRPPRAGLLSAGSSGAGSRLQAEMLQMDLVDAAGDTPCAEDDEEDDELLAA 119
QY 120 RRPAGAPPAESGOGEPASRGQSGQSGQSGQSGQSGQSGQSGQSGQSGQSGQSGQSG 179
Db 120 QRPVGPPKAESNODPAPR-----SQGGQSGTSGDTYRPRRTTLLNLFQVPRSDTLNN 175
QY 180 NSLGKHSWQDRVSRSSSPLKTGEOTPPHEHICLSBELPPQSGPAPTTDRGTSTDSPCR 239
Db 176 NSLGKHSWQDRVSRSSSPLKTGEOTPPHEHICLSBELPPQSGPAPTTDRGTSTDSPCR 235
QY 240 STATOMAPPGPPAAPPGGGRGSHRDIHYQADVRLAEATEEYILTPVQRPDPAEPTSAF 299
Db 236 SAATOMAPPGPPAPGGRGSHRDIHYQADVRLAEATEEYILTPVQRPDPAEPTSTF 295
QY 300 LPPTESRMSVSDPDPAAYPSTAGRPHPSISEEEGFDCLSSPERAEPGGGWRGSLGEP 359
Db 296 MPPTESRMSVSDPDPAAYSVTAGRPHPSISEEEGFDCLSSPERAEPGGGWRGSLGEP 355
QY 360 PPPRASLSSTALSYSYKTYLVDEHAQLVSLRCPGQSGYSDSDSATVYDNCASV 419
Db 356 PPPRASLSSTALSYSYKTYLVDEHAQLVSLRCPGQSGYSDSDSATVYDNCASA 415
QY 420 SSPYESAIGEYEEAPRQPPACLSDESTDPEDVHFSSKFLNFMVMSGRSSSSAESFGL 479
Db 416 SSPYESAIGEYEEAPRQPPACLSDESTDPEDVHFSSKFLNFMVMSGRSSSSAESFGL 475
QY 480 FSCIINGEEOETHRAIFRFVPRHEDELEVDPLLVQLAEDWYKAYNRTGARGVF 539
Db 476 FSCVINGEEOETHRAIFRFVPRHEDELEVDPLLVQLAEDWYKAYNRTGARGVF 535
QY 540 PAYYAEVTKPEHMAALAKNSDWVQFRVYKFLGVSQVPYHKGNDVLCAMQKIATTRL 599
Db 536 PAYYAEVTKPEHMAALAKN----- 556
QY 600 TVHFNPPSSCVLEISVRGKIGVKADDSOBKGNKCSHFQLNISFCGYHPKNNKYFGF 659
Db 557 -----SCVLEISVRGKIGVKADDALEAKGNKCSHFQLNISFCGYHPKNNKYFGF 608
QY 660 ITKHPADHRFACHVFVEDSTKALAESVGRATFOFYKQFVEYTCPTEDIYLE 711
Db 609 ITKHPADHRFACHVFVEDSTKALAESVGRATFOFYKQFVEYTCPTEDIYLE 660

RESULT 9
AAG78824
ID AAG78824 standard; Protein; 673 AA.
XX
AC AAG78824;
XX
DT 18-DEC-2001 (first entry)
XX
DE Murine SKIP-2b.
XX
KW Murine; SKIP-2b; SAPK-interacting protein; phosphorylase; JNK; SAPK;
KW c-Jun N-terminal kinase; Stress-Activated Protein Kinase;
XX cell death related disease.
XX
OS Mus musculus.

```

```

XX PN KR2001029352-A.
XX PD 06-APR-2001.
XX PF 30-SEP-1999; 99KR-0042118.
XX PR 30-SEP-1999; 99KR-0042118.
XX PA (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.
XX PI Choi IY, Choi UY, Han PR, Lee GU, Lee JG, Lee SH;
XX DR WPI; 2001-563617/63.
XX DR N-PSDB; AAI65089.
XX PT Mouse skip (SAPK-interacting protein) controlling activation of human
XX PR cell phosphorylase, jnk, and its gene, useful for treating cell death
XX PS related diseases.
XX PS Claim 2; Page 12-15; 37pp; Korean.
XX CC The present sequence is the protein sequence for murine SKIP-2b
XX CC (SAPK-interacting protein), which selectively inhibits or controls the
XX CC activation of a phosphorylase, JNK/SAPK (c-Jun N-terminal kinase/
XX CC Stress-Activated Protein Kinase), activated by various kinds of stress.
XX CC SKIP and its gene are useful in treating cell death related diseases.
XX SQ Sequence 673 AA;

Query Match 85.7%; Score 3248; DB 22; Length 673;
Best Local Similarity 90.0%; Pred. No. 1.8e-245;
Matches 610; Conservative 13; Mismatches 25; Indels 30; Gaps 3;

QY 35 LTHDISLEFEDEDESEITDCGSLQCKDTLSLRPRAGLLSAGGAGSRLQAFMLQ 94
DB 25 LTHDISLEFEDEDESEITDCGSLQCKDTLSLR-----M 60

QY 95 DLIDATGDTFGAEDDEDDDEARPRGAGPPKAESGQEFASRGQSGQSGQSGSD 154
DB 61 DLIDAAAGDTFGAEDDEDEDEDELAARPGVGPKAESNQDPAPRSQGGQ-----ATGSGD 115

QY 155 TYPRKPTTLNLFQVPRSQDTLNNSLGKHSQDRVSRSSPLKTGTQTPPHEHICLS 214
DB 116 TYPRKPTTLNLFQVPRSQDTLNNSLGKHSQDRVSRSSPLKTGTQTPPHEHICLS 175

QY 215 EELPQSGPAPTTDRGTSTDSFCRRSTATQMAPPGPPAAPPCCRGHSHRDRHIQADV 274
DB 176 DELPPQSGVPVTDRTSTDSFCRRSAATQMAPPGPPAAPPCCRGHSHRDRHIQADV 235

QY 275 LEATEEIIYLPVQRPDAAEPTSAFLPPTESRMSVSSDDPDFAAYSTAGRPHPSISEE 334
DB 236 LEATEEIIYLPVQRPDPAETSTFMPPTESRMSVSSDDPDFAAYSTAGRPHPSISEE 295

QY 335 GFDCLSLSPERAPPGGWRGSLGEPDPPPRASLSSDTSALSYDVSKYTLVDEHAQLEIV 394
DB 296 GFDCLSLSPERAPPGGWRGSLGEPDPPPRASLSSDTSALSYDVSKYTLVDEHAQLEIV 355

QY 395 SLRPGCDYSDSDSATYDNCASVSPYESAIGEYEEAPRQPPACLSSESTDPEDPV 454
DB 356 SLRPGCDYSDSDSATYDNCASVSPYESAIGEYEEAPRQPPACLSSESTDPEDPV 415

QY 455 HFSKKFLNFMGSRSSSSAESSFLGSCINGEEQTHRAIFRFPVPRHEDELELEVDPP 514
DB 416 HFSKKFLNFMGSRSSSSAESSFLGSCINGEEQTHRAIFRFPVPRHEDELELEVDPP 475

QY 515 LLVELQAEDYWEAYNMTGARGVFPAYATVETKPEHMAALAKNSDMVDQFRVKFLGS 574
DB 476 LLVELQAEDYWEAYNMTGARGVFPAYATVETKPEHMAALAKNSDMVDQFRVKFLGS 535

QY 575 VQVPYHKGNDVLCAMQKIATR-RLNVHNPSSCVLETISRGVKGIVKADDSQAKGN 633
DB 536 VQVPYHKGNDVLCAMQKIATRRLNRPRTLVHNPSSCVLETISRGVKGIVKADDALEAKGN 595

```

```

QY 634 KCSHFFOLKNISFCGYHPKNNKYFGFTTKHPADHRFACHVFSVSEDSTKALAESVGRAFQ 693
DB 596 KCSHFFOLKNISFCGYHPKNNKYFGFTTKHPADHRFACHVFSVSEDSTKALAESVGRAFQ 655

QY 694 FYKQFVEVTCPTEDIYLE 711
DB 656 FYKQFVEVTCPTEDIYLE 673

RESULT 10
ABB04788
ID ABB04788 standard; Protein; 659 AA.
XX AC ABB04788;
XX DT 13-MAR-2002 (first entry)
XX DE LDL receptor binding protein JIP-1 SEQ ID NO:11.
XX KW Low density lipoprotein receptor binding protein; signal transduction;
XX KW LDL receptor binding protein; LDL receptor signalling pathway.
XX OS Synthetic.
XX PN WO200184159-A2.
XX PD 08-NOV-2001.
XX PF 24-APR-2001; 2001WO-US13214.
XX PR 01-MAY-2000; 2000US-0562737.
XX PA (TEXA) UNIV TEXAS SYSTEM.
XX PI Herz J, Gotthardt M;
XX WPI; 2002-082855/11.

CC Detecting stress that alters interaction of LDL receptor binding
CC polypeptide with LDL receptor interaction domain, comprises detecting
CC difference in stress-biased and unbiased interaction of peptide and
CC domain in a system.
CC Disclosure; Page 34-36; 200pp; English.

CC The present invention describes a method for detecting a stress that
CC alters a functional interaction of a low density lipoprotein (LDL)
CC receptor binding protein (I) with an LDL receptor interaction domain
CC (II). The method involves introducing a predetermined stress into a
CC system which provides a stress-biased physical interaction of (I) with
CC (II), where in the absence of the stress, the system provides an
CC unbiased interaction of (I) and (II), and detecting the stress-biased
CC interaction of (I) and (II), where a difference between BI and UI
CC indicates that the stress alters the interaction of (I) and (II). (I)
CC is selected from SEMCAP-1, JIP-1, PSD-95, JIP-2, talin, OMP25, CAPON,
CC p14,5 Kinase, Na channel brain 3, Mint1, ICAP-1 and APC subunit 10.
CC The method is useful for detecting a stress that alters functional
CC interaction of LDL receptor binding polypeptide with LDL receptor
CC interaction domain. The method is useful for detecting and modulating
CC signal transduction through LDL receptors. ABB04778 to ABB04909
CC represent LDL receptor binding proteins which are used in the
CC exemplification of the present invention.
XX SQ Sequence 659 AA;

```

```

Query Match 85.6%; Score 3245; DB 23; Length 659;
Best Local Similarity 87.1%; Pred. No. 2.9e-245;
Matches 620; Conservative 12; Mismatches 26; Indels 54; Gaps 5;

QY 1 MAERSSGLGGGAASPPAASPFLGLHTASPPNRLTHDLSLEFEDEDESEITDCGSL 60
DB 1 MAERSS-GLGGGAASPPAASPFLGLHTASPPNRLTHDLSLEFEDEDESEITDCGSL 59

```



```
QY 61 QCKDTLSLRPPRAGLLSAG-GGGAGSRLOAQLMDLIDATGDTPGAEDDEDDERAA 119
Db 60 QCKDTLSLRPPRAGLLSAGSSGSRLOAQLMDLIDAGDTPGAEDDEDEELAA 119
QY 120 RPPGAGPPKAESQGPASRGOSQSGQSGGDTYRKRPTTLNLFPPQVPRSDTLNN 179
Db 120 QRPVGPPKAESNQDPAPR-SSQOGPGTSGDTPYRKRPTTLNLFPPQVPRSDTLNN 175
QY 180 NSLGGKHSWQDRVSRSSPLKGTGQTPHEHICLSEELPPOSGPAPTTDRGTSTDSPCRR 239
Db 176 NSLGGKHSWQDRVSRSSPLKGTGQTPHEHICLSEELPPOSGPAPTTDRGTSTDSPCRR 235
QY 240 STATQMAPGGPAPPGGGRGSHRDRIHQADVLEATEEITLTPVQRPDAEPTSAF 299
Db 236 SAATQMAPGGPAPPGGGRGSHRDRIHQADVLEATEEITLTPVQRPDAEPTSAF 295
QY 300 LPPTSRMSVSSDDPPAAYPSTAGRPSPSISEEERGFDCLSPERAEPGGGWRGSLGEP 359
Db 296 MPPTSRMSVSSDDPPAAYSVTAGRPSPSISEEERGFDCLSPERAEPGGGWRGSLGEP 355
QY 360 PPPPRASLSSTALSYSVSKYTLVVDHAQLVSLRCPFGDYSDESATSATVYDNCASV 419
Db 356 PPPPRASLSSTALSYSVSKYTLVVDHAQLVSLRCPFGDYSDESATSATVYDNCASA 415
QY 420 SSPYESAIGEEYEAAPRPPQACLSDETPDEPDVHFSKKFLNFMGSRSSSAESFGL 479
Db 416 SSPYESAIGEEYEAAPRPPQACLSDETPDEPDV-FSKKFLNFMGSRSSSAESFGL 474
QY 480 FSCIINGEQEQTHRAIFRFRVPRHEDELEVDLPLVELQAEDYWEAYNMRTGARGVF 539
Db 475 FSCVINGEEHQTHRAIFRFRVPRHEDELEVDLPLVELQAEDYWEAYNMRTGARGVF 534
QY 540 PYYAIEVTKPEPHEMALAKNSDWQDFRVKFLGVSQVYHKGNDVLCAMQKIATTRRL 599
Db 535 PYYAIEVTKPEPHEMALAKN----- 555
QY 600 TVHFNPPSSCVLEISVRGVKIGVADDSQEAQGNKCSHFOLKNISFCGYHPKNNKYFGF 659
Db 556 -----SCVLEISVRGVKIGVADDALEAKGNKCSHFOLKNISFCGYHPKNNKYFGF 607
QY 660 ITKHPADHRFACHVFVSDSTKALAESVGRFAQFYKQFVEYTCPTEDIYLE 711
Db 608 ITKHPADHRFACHVFVSDSTKALAESVGRFAQFYKQFVEYTCPTEDIYLE 659

RESULT 11
AAG78827
ID AAG78827 standard; Protein; 617 AA.
XX
AC AAG78827;
XX
DT 18-DEC-2001 (first entry)
XX
DE Murine SKIP-3.
XX
KW Murine; SKIP-3; SAPK-interacting protein; phosphorylase; JNK; SAPK;
KW c-Jun N-terminal kinase; Stress-Activated Protein Kinase;
KW cell death related disease.
XX
OS Mus musculus.
XX
PN KR2001029352-A.
XX
PD 06-APR-2001.
XX
PF 30-SEP-1999; 99KR-0042118.
XX
PR 30-SEP-1999; 99KR-0042118.
XX
PA (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.
XX
PI Choi IY, Choi UJ, Han PR, Lee GU, Lee JG, Lee SH;
```

```
XX
DR N-PSDB; AAI65092.
XX
PT Mouse skip (SAPK-interacting protein) controlling activation of human
PT cell phosphorylase, jnk, and its gene, useful for treating cell death
PT related diseases -
XX
PS Claim 2; Page 30-33; 37pp; Korean.
XX
CC The present sequence is the protein sequence for murine SKIP-3
CC (SAPK-interacting protein), which selectively inhibits or controls the
CC activation of a phosphorylase, JNK/SAPK (c-Jun N-terminal kinase/
CC Stress-Activated Protein Kinase), activated by various kinds of stress.
CC SKIP and its gene are useful in treating cell death related diseases.
XX
SQ Sequence 617 AA;

Query Match 81.9%; Score 3104; DB 22; Length 617;
Best Local Similarity 92.9%; Pred. No. 2.9e-234;
Matches 578; Conservative 13; Mismatches 25; Indels 6; Gaps 2;

QY 91 MLOWDLIDATGDTPGAEDDEDDERARRPCGAPPKAESGOEPASRGOSQSGQP 150
Db 1 MLOWDLIDAGDTPGAEDDEDEELAAQRCVGPVKAESNQDPAPRSGQG-AT 55
QY 151 GSGDTYRKRPTTLNLFPOVPRSDTLNNSLGKHSWQDRVSRSSPLKGTGQTPPHEH 210
Db 56 GSGDTYRKRPTTLNLFPOVPRSDTLNNSLGKHSWQDRVSRSSPLKGTGQTPPHEH 115
QY 211 ICUSEELPQSGPAPTDRGTSTDSPCRRSTATQMAPPGGPPAAGGGRGSHRDRIHQ 270
Db 116 ICULDELPPQSGVPVQDRGTSTDSPCRRSAATQMAPPGPPATAPGGRGSHRDRIHQ 175
QY 271 ADVRLATEEITLTPVQRPDAEPTSAFLPTESRMSVSSDDPPAAYSTAGRPSPSIS 330
Db 176 ADVRLATEEITLTPVQRPDAEPTSTFMPPTESRMSVSSDDPPAAYSTAGRPSPSIS 235
QY 331 EEEEGFDCLSLSPERAPPGGWRGSLGEPPPPRASSLSDTSALSYDSVKYTLVVDHAQ 390
Db 236 EEEEGFDCLSLSPERAPPGGWRGSLGEPPPPRASSLSDTSALSYDSVKYTLVVDHAQ 295
QY 391 LELVSLRCPFGDYSDESATSATVYDNCASVSPYESAIGEEYEAAPRPPQACLSDETPD 450
Db 296 LELVSLRCPFGDYSDESATSATVYDNCASVSPYESAIGEEYEAAPRPPQACLSDETPD 355
QY 451 EPDVHFSKKFLNFMGSRSSSAESFGLFSCIINGEQEQTHRAIFRVPVRHEDELE 510
Db 356 EPDVHFSKKFLNFMGSRSSSAESFGLFSCIINGEQEQTHRAIFRVPVRHEDELE 415
QY 511 VDDPLLVELQAEDYWEAYNMRTGARGVFPAYYAEVTKPEPHEMALAKNSDWQDFRVK 570
Db 416 VDDPLLVELQAEDYWEAYNMRTGARGVFPAYYAEVTKPEPHEMALAKNSDWQDFRVK 475
QY 571 FLGVSQVYHKGNDVLCAMQKIATTR-RLTVHFNPPSSCVLEISVRGVKIGVADDSQ 629
Db 476 FLGVSQVYHKGNDVLCAMQKIATTRPRLTVHFNPPSSCVLEISVRGVKIGVADDALE 535
QY 630 AKGNKCSHFOLKNISFCGYHPKNNKYFGFITKHPADHRFACHVFVSDSTKALAESVGR 689
Db 536 AKGNKCSHFOLKNISFCGYHPKNNKYFGFITKHPADHRFACHVFVSDSTKALAESVGR 595
QY 690 AFQFYKQFVEYTCPTEDIYLE 711
Db 596 AFQFYKQFVEYTCPTEDIYLE 617

RESULT 12
ABB04797
ID ABB04797 standard; Protein; 659 AA.
XX
AC ABB04797;
XX
```

DT 13-MAR-2002 (first entry)  
XX LDL receptor binding protein JIP-1 SEQ ID NO:20.  
DE Low density lipoprotein receptor binding protein; signal transduction;  
XX LDL receptor binding protein; LDL receptor signalling pathway.  
KW Synthetic.  
XX WO200184159-A2.  
PN 08-NOV-2001.  
XX 24-APR-2001; 2001WO-US13214.  
PF 01-MAY-2000; 2000US-0562737.  
XX (TEXA ) UNIV TEXAS SYSTEM.  
PA Herz J, Gotthardt M;  
XX WPI; 2002-082855/11.  
PI Detecting stress that alters interaction of LDL receptor binding  
XX polypeptide with LDL receptor interaction domain, comprises detecting  
XX difference in stress-biased and unbiased interaction of peptide and  
XX domain in a system -  
PS Disclosure; Page 49-51; 200pp; English.  
XX The present invention describes a method for detecting a stress that  
XX alters a functional interaction of a low density lipoprotein (LDL)  
XX receptor binding protein (I) with an LDL receptor interaction domain  
XX (II). The method involves introducing a predetermined stress into a  
XX system which provides a stress-biased physical interaction of (I) with  
XX (II), where in the absence of the stress, the system provides an  
XX unbiased interaction of (I) and (II), and detecting the stress-biased  
XX interaction of (I) and (II), where a difference between BI and UI  
XX indicates that the stress alters the interaction of (I) and (II). (I)  
XX is selected from SEMCAP-1, JIP-1, PSD-95, JIP-2, Talin, OMP25, CAPON,  
XX PI4,5 Kinase, Na channel brain 3, Mint1, ICAP-1 and APC subunit 10.  
XX The method is useful for detecting a stress that alters functional  
XX interaction of LDL receptor binding polypeptide with LDL receptor  
XX interaction domain. The method is useful for detecting and modulating  
XX signal transduction through LDL receptors. ABB04778 to ABB04909  
XX represent LDL receptor binding proteins which are used in the  
XX exemplification of the present invention.  
XX Sequence 659 AA;  
Query Match 75.8%; Score 2875; DB 23; Length 659;  
Best Local Similarity 78.8%; Pred. No. 2.7e-216;  
Matches 561; Conservative 18; Mismatches 79; Indels 54; Gaps 5;  
QY 1 MAERESGLGGGAAPPAAPSPFLGLHIAAPNPRFLTHDISLEEFDEDELSITDECGISL 60  
DB 1 MAERES-GLGAGAAAPPAASDFLGLHIAAPNPRFLTHDISFEFEDEDELSITDECGISL 59  
QY 61 QCKDTLSLRPRRAGLSAGG-CGAGSRLQAMQLMDLIDATGTPGAEDEDEDEERA 119  
DB 60 HCKDTLSLRIRAGLSAGSGSAGSRLQALMDLIDAMGDTGAEADNEEDDELA 119  
QY 120 RRGAGPPKAESQEPASRCOGSOGSOGSGDTRPKRPTLNLFPQVPRSDTLNN 179  
DB 120 RRGVGGPKASSNODPAPR----STGGGTGSGVTRPKRPTLNLFPQVPRSDTLNN 175  
QY 180 NSLCKHSHWDVRSRSPSLKTEQTPPHHICLSBELPPQSGPAPTDRGTSTDSPCR 239  
DB 176 NSLGAKHSHWDVRSRSPSLKTEQTPPHHICLSBELPPQSGPAPTDRGTSTDSPCR 235  
QY 240 STAFOMAPGGPPAAPPGGRGHSHRDRIHQADVRLAEATEEYILTPVQRPDAEPTSAF 299  
DB 236 SAATIMAPSGPPAKAPGGRGHSHLDRIHQADVRLAEATEEYILNPVQRPDAEPTSTF 295

QY 300 LPPTESRMSSVSDPDPAAYPSTAGRPHPSISEEEGFDCLSSPERAEPGGWGRSLGEP 359  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 355  
QY 360 PPPRASLSDSSTALSYSVKYTLVVDHAQLELVLRLPCFGDYSDSDSATVYDNCASV 419  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 415  
QY 420 SSPYESAIGEEYEAPRPPACLSDESDTDPDVFHFSKKFLNFMVMSGRSSSAESFGL 479  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 474  
QY 480 FSCITNGEQEOTHRALFRFVPRHDELEVDLPLVQLQAEYDYWEAYNNRTGARGVF 539  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 534  
QY 540 PAYYAIETVKEPEHMAALAKNSDWQDFRYKFLGVSQVPHYHKGNDVLCAAMOKIATTRL 599  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 555  
QY 600 TVHFNPPSSCVLEISVRGKIGVKADDSQAKNCKSHFPOLKNISFCGYHPKNNKYFGF 659  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 607  
QY 660 ITKHPADHREACHVFVSEDSSTKALAESVGRFAQQFYKQFVEYTCPTEDIYLE 711  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 659  
RESULT 13  
ABB04793  
ID ABB04793 standard; Protein; 659 AA.  
XX ABB04793;  
XX AC  
XX 13-MAR-2002 (first entry)  
XX DT  
XX LDL receptor binding protein JIP-1 SEQ ID NO:16.  
XX DE  
XX Low density lipoprotein receptor binding protein; signal transduction;  
KW LDL receptor binding protein; LDL receptor signalling pathway.  
XX OS Synthetic.  
XX PN WO200184159-A2.  
XX PD 08-NOV-2001.  
XX 24-APR-2001; 2001WO-US13214.  
XX PF 01-MAY-2000; 2000US-0562737.  
XX PR (TEXA ) UNIV TEXAS SYSTEM.  
XX PA Herz J, Gotthardt M;  
XX PI WPI; 2002-082855/11.  
XX DR  
XX Detecting stress that alters interaction of LDL receptor binding  
PT polypeptide with LDL receptor interaction domain, comprises detecting  
PT difference in stress-biased and unbiased interaction of peptide and  
PT domain in a system -  
XX PS Disclosure; Page 43-44; 200pp; English.  
XX The present invention describes a method for detecting a stress that  
XX alters a functional interaction of a low density lipoprotein (LDL)  
XX receptor binding protein (I) with an LDL receptor interaction domain  
XX (II). The method involves introducing a predetermined stress into a  
XX system which provides a stress-biased physical interaction of (I) with  
XX (II), where in the absence of the stress, the system provides an  
XX unbiased interaction of (I) and (II), and detecting the stress-biased

```
CC interaction of (I) and (II), where a difference between BI and UI
CC indicates that the stress alters the interaction of (I) and (II). (I
CC is selected from SEMCAP-1, JIP-1, PSD-95, JIP-2, Talin, OMP25, CAPON,
CC PIP4,5 Kinase, Na channel brain 3, Mint1, ICAP-1 and APC subunit 10.
CC The method is useful for detecting a stress that alters functional
CC interaction of LDL receptor binding polypeptide with LDL receptor
CC interaction domain. The method is useful for detecting and modulating
CC signal transduction through LDL receptors. ABB04778 to ABB04909
CC represent LDL receptor binding proteins which are used in the
CC exemplification of the present invention.
XX
SQ Sequence 659 AA;

Query Match 75.5%; Score 2864; DB 23; Length 659;
Best Local Similarity 78.5%; Pred. No. 1.9e-215;
Matches 559; Conservative 23; Mismatches 76; Indels 54; Gaps 5;

Qy 1 MAERESGLGGGAASPPAASPFGLGHIAAPPNFRFLTHDISLEEFEDLSEITDECGISL 60
Db 1 MAEREA-GLGGGAASPPAASPFGLGHIAAPPNFRFLTHDISLEEFEDGLSEITDECHISL 59

Qy 61 QCKDTLSLRPRAGLLSAG-GGAGSRLQAEMLQMDLIDATGTPCAEDDEDDDEERAA 119
Db 60 QCKDTLSLRPRAGLLKAGSGSAGSLQAEMLQMDIDAAGTPEGNEDEDEEDQLAA 119

Qy 120 RPPGAGPPKAESGOEPASRGOGSGOGSGDTPYRKRPPTLNLFPQVPRSQDTLNN 179
Db 120 QRPVGVRPKAESNDPSPR----SOGOGVGSGDTPYRKRPPTLNLFPQVPRSQDTLNN 175

Qy 180 NSLGKHSWQDRVSRSSPLKTGEQTPPHEHICLSELPPQSGPAPTTDRGTSTDSPCR 239
Db 176 ASLGKHSWQDRVSRSSPLKTGEQTPPHEHICLSELPPQSGPAPTTDRGTSTDSPCR 235

Qy 240 STATOMAPPGPPAAPGGGRGSHRDRHYQADVRLATEEIIYLTVPQRPDAEPTSAF 299
Db 236 IAATQMAPPKPPATAPGRLSHRDRHYMADVRLATEENIYLTVPQRPDAEPTSTF 295

Qy 300 LPPTESRMVSSDDPPAAYSTAGRPHPSISEEEGFDCLSSPERAEPGCGWRGSLGEP 359
Db 296 RPPTESRMVTSDDPPAAYSWTAGRPHPSIYEDEGFDCLASPERAEPGCGWRGSLGEP 355

Qy 360 PPPRASLSSDTSALSVDYKTYTLVDEHAQLVLRPCFGDYSDSDSATYVDNCASV 419
Db 356 EPPRASLSSDTSALSVDYKTYTLVDEHAQLVLRPCIGDYSDSDSKTYVDNCASA 415

Qy 420 SSPYSAIGEYEEAPRPPACLSDESTDPDPVHFSKKFLNVFMSGRSSSAESFGL 479
Db 416 LSPYSAIGEYEEAPRPPACLSDESTDPDPV-FSKKERNVFMGRSRTSSAESFGL 474

Qy 480 FSCIINGEQQTHRAIFRFVPRHEDELELVDDPLLVQLAEADYWEAYNMRGTGARGVF 539
Db 475 FVCVINGEEHWHTRAIFRFVPRHEDELEADPLLVLEQDEYWEAYNERTGARGVF 534

Qy 540 PAYVAIVTEKPEHMAALAKNSWDQFRVKFLGSGVQVYHKGNVDVLCAMQKTIATTRL 599
Db 535 PFYVAIVTEKGEHMAALAKN-----

Qy 600 TVHENPPSSCVLETSVRGKTVKADDSOEAQGNKCSHFQLKNI SPCGVKPNKNKYFGF 659
Db 556 -----HCVLEISVVRGIKIGVKADDAKEAGNKCASHLQKNI SFCMTYHPRNKNYFN 607

Qy 660 ITKHPADHRFACHVFSVSDTSKALAESVGRFAFOQFYKQFVEYTCPTEDIYLE 711
Db 608 ITKHPADHQFACHVFSVSRSTKALAESVSRFAQFYKQFVEYTCPTEDVYLE 659

RESULT 14
ABB04790
ID ABB04790 standard; Protein: 659 AA.
XX
AC ABB04790;
XX
DT 13-MAR-2002 (first entry)
```

```
XX LDL receptor binding protein JIP-1 SEQ ID NO:13.
DE
XX Low density lipoprotein receptor binding protein; signal transduction;
KW LDL receptor binding protein; LDL receptor signalling pathway.
XX Synthetic.
XX WO200184159-A2.
XX 08-NOV-2001.
XX 24-APR-2001; 2001WO-US13214.
XX 01-MAY-2000; 2000US-0562737.
XX (TEXA ) UNIV TEXAS SYSTEM.
XX Herz J, Gotthardt M;
XX WPI; 2002-082855/11.
XX Detecting stress that alters interaction of LDL receptor binding
PT polypeptide with LDL receptor interaction domain, comprises detecting
PT difference in stress-biased and unbiased interaction of peptide and
PT domain in a system -
XX
PS Disclosure; Page 38-39; 200pp; English.
XX
CC The present invention describes a method for detecting a stress that
CC alters a functional interaction of a low density lipoprotein (LDL)
CC receptor binding protein (I) with an LDL receptor interaction domain
CC (II). The method involves introducing a predetermined stress into a
CC system which provides a stress-biased physical interaction of (I) with
CC (II), where in the absence of the stress, the system provides an
CC unbiased interaction of (I) and (II), and detecting the stress-biased
CC interaction of (I) and (II), where a difference between BI and UI
CC indicates that the stress alters the interaction of (I) and (II). (I
CC is selected from SEMCAP-1, JIP-1, PSD-95, JIP-2, Talin, OMP25, CAPON,
CC PIP4,5 Kinase, Na channel brain 3, Mint1, ICAP-1 and APC subunit 10.
CC The method is useful for detecting a stress that alters functional
CC interaction of LDL receptor binding polypeptide with LDL receptor
CC interaction domain. The method is useful for detecting and modulating
CC signal transduction through LDL receptors. ABB04778 to ABB04909
CC represent LDL receptor binding proteins which are used in the
CC exemplification of the present invention.
XX
SQ Sequence 659 AA;

Query Match 75.1%; Score 2846; DB 23; Length 659;
Best Local Similarity 78.4%; Pred. No. 5e-214;
Matches 558; Conservative 20; Mismatches 80; Indels 54; Gaps 5;

Qy 1 MAERESGLGGGAASPPAASPFGLGHIAAPPNFRFLTHDISLEEFEDLSEITDECGISL 60
Db 1 MAARES-GLUGGADSPPAASPFLELGHIAAPPNFFLTHDISLEEGEDELSEITHCGISL 59

Qy 61 QCKDTLSLRPRAGLLSAGGGA-GSRLQAEMLQMDLIDATGTPCAEDDEDDDEERAA 119
Db 60 QCKTILSLRPPRAKLLSAGSSGLSRLQAEMLMDLIDAGDNPGEAEDDEEQDDELA 119

Qy 120 RPPGAGPPKAESGOEPASRGOGSGOGSGDTPYRKRPPTLNLFPQVPRSQDTLNN 179
Db 120 QRPVRVPPKAESNSDPAPR---SQGTPTGSGDTPYRKRPPTLNLMPQVPRSQDTYNN 175

Qy 180 NSLGKHSWQDRVSRSSPLKTGEQTPPHEHICLSELPPQSGPAPTTDRGTSTDSPCR 239
Db 176 NSLGKHWQDRVSRSSDLKTEQTPPEHICLSLSEDEFPPQSGSPVPTGRTGRTSDSPHR 235

Qy 240 STATOMAPPGPPAAPGGGRGSHRDRHYQADVRLATEEIIYLTVPQRPDAEPTSAF 299
Db 236 SAATOMAI SGPPTAPKGRGSHRDRHYQADVRLATEEIIYLTVPNRPDPAPETQTF 295
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2002, 08:48:50 ; Search time 26 Seconds  
(without alignments)  
2628.907 Million cell updates/sec

Title: US-09-966-561-2

Perfect score: 3792

Sequence: 1 MAERESGLGGGAASPPAAS.....QQFYKQFVEYTCPTEDIYLE 711

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR73: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3263.5	86.1	660	2 T03038	probable inhibitor
2	2463.5	6.5	315	2 A86043	protein C13A10.3 [
3	195.5	5.2	705	2 A5363	synapsin I splice
4	194.5	5.1	2142	2 B35098	MHC class III hist
5	191	5.0	1872	2 S36152	MHC class III hist
6	190.5	5.0	1870	2 S37671	MHC class III hist
7	184.5	4.9	706	2 E30411	synapsin Ia - bovi
8	183	4.8	1733	2 S27939	synapsin - chicken
9	182	4.8	1634	2 T26517	hypothetical prote
10	179.5	4.7	4957	2 T03455	ALR protein - huma
11	179.5	4.7	5262	2 T03454	ALR protein - huma
12	179	4.7	1744	2 A54970	tensin, cardiac mu
13	179	4.7	1792	2 A57075	tensin - chicken (
14	177.5	4.7	1240	2 J55209	insulin receptor s
15	177.5	4.7	1242	2 J50670	insulin receptor s
16	174.5	4.6	1077	2 A44067	serine-rich protei
17	174	4.6	1184	2 G01763	atrophin-1 - huma
18	173.5	4.6	403	2 S22796	hypothetical prote
19	173.5	4.6	1110	2 T19673	prpL2 protein - hu
20	171.5	4.5	1279	2 T18312	hypothetical prote
21	170.5	4.5	440	2 J57807	Wiskott-Aldrich sy
22	170.5	4.5	1664	2 T18262	S-layer protein -
23	170	4.5	3149	1 Q08E8	BPLF1 protein - hu
24	169.5	4.5	1184	2 S00832	atrophin-1 - huma
25	169	4.5	1231	2 S30185	insulin receptor s
26	167.5	4.4	1958	2 B40505	hypothetical prote
27	167	4.4	1262	2 T13353	protein stn-B - fr
28	166.5	4.4	1235	1 S16948	insulin receptor s
29	166	4.4	704	2 A30411	synapsin Ia - rat

## ALIGNMENTS

### RESULT 1

T03038  
probable inhibitor protein JIP-1 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 08-Oct-1999  
C:Accession: T03038  
R:Dickens, M.; Rogers, J.S.; Cavanagh, J.; Raitano, A.; Xia, Z.; Halpern, J.R.; Green  
Science 277, 693-696, 1997  
A:Title: A cytoplasmic inhibitor of the JNK signal transduction pathway.  
A:Reference number: 214833; MUID:97382313; PMID:9235893  
A:Accession: T03038  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-660 <DIC>  
A:Cross-References: EMBL:AF003115; NID:g2316073; PIDN:AAB66317.1; PID:g2316074  
C:Genetics:  
A:Note: JIP-1  
C:Function:  
A:Description: binds to the JNK protein kinase and inhibits JNK signal transduction p

Query Match	86.1%	Score	3263.5	DB	2	Length	660
Best Local Similarity	87.4%	Pred. No.	2.6e-175				
Matches	621	Conservative	12	Mismatches	26	Indels	53
Gaps	4						
QY	1	MAERESGLGGGAASPPAASPFGLGHIAFPNPRHLTHDISEEFDEDLSEITDECGISL	60				
DB	1	MAERES-GLGGGAASPPAASPFGLGHIAFPNPRHLTHDISEEFDEDLSEITDECGISL	59				
QY	61	QCKDTLSLRPPRAGLLSAG-GGGAGSRQAEMQLQMDLIDATGDTFGAEDDEEDDEERAA	119				
DB	60	QCKDTLSLRPPRAGLLSAGSGSAGSRQAEMQLQMDLIDAGDTFGAEDDEEDDEELAA	119				
QY	120	RRPGAGPKAESGGEPASRGSGOGSGOGSGDYRPRKPTTLNLFQVPRSQDTLNN	179				
DB	120	RRPGVGPKAESNODPAPR---SQGGPGTSGDYRPRKPTTLNLFQVPRSQDTLNN	175				
QY	180	NSLGHKHHNQDVRSSRSPDKTGEQTPPHHICLSEELPPQSGPAPTTDRTSTDSCCR	239				
DB	176	NSLGHKHHNQDVRSSRSPDKTGEQTPPHHICLSEELPPQSGPAPTTDRTSTDSCCR	235				
QY	240	STATQMAPGGPPGAPPGHSHRDRHYQADVRLAEATEEYILTPVQRPDAAEPTSAF	299				
DB	236	SAATQMAPGGPPGAPPGHSHRDRHYQADVRLAEATEEYILTPVQRPDAAEPTSAF	295				
QY	300	LPTETSRMSVSDPDPAAYSTAGRPHSPSTSEEGFDCLSGPERAEPGGWGSGLGE	359				
DB	296	MPTETSRMSVSDPDPAAYSTAGRPHSPSTSEEGFDCLSGPERAEPGGWGSGLGE	355				
QY	360	PPPRASLSSTPSALSVDYVKYTLVDEHAQLELYSLRPFCDGYSDESATVYDNCASV	419				
DB	356	PPPRASLSSTPSALSVDYVKYTLVDEHAQLELYSLRPFCDGYSDESATVYDNCASA	415				

```

QY 420 SSPYSAIGEBYBAPRQPPACLSDESDPDEVDHFSKFLNVPMSGRSSSAESFGL 479
Db 416 SSPYSAIGEBYBAPRQPPACLSDESDPDEVDHFSKFLNVPMSGRSSSAESFGL 475
QY 480 FSCIINGEQTHRAIFRVPVRHEDELEVDPLVQLQAEYDYWEAYNNRTGARGVF 539
Db 476 FSCVINGEHEQTHRAIFRVPVRHEDELEVDPLVQLQAEYDYWEAYNNRTGARGVF 535
QY 540 PAYIAIEVTKPEHMAALAKNSDWDFRKFGLGVSQVDPYHKGNDVLCAAMQKATATTRL 599
Db 536 PAYIAIEVTKPEHMAALAKN----- 556
QY 600 TVHFNPSSCVLEISVRGKIGVAKADDSQPAKNGKSHFPQKNISFCGYHPKNNKYFGF 659
Db 557 -----SCVLEISVRGKIGVAKADDALEAKNGKSHFPQKNISFCGYHPKNNKYFGF 608
QY 660 ITKHPADHRFACHVFSVSEDSTKALAESVGRFAFOFYKQFVEYTCPTEDIYLE 711
Db 609 ITKHPADHRFACHVFSVSEDSTKALAESVGRFAFOFYKQFVEYTCPTEDIYLE 660

RESULT 2
A88043
protein C13A10.3 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: A88043
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/asc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: A88043
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-315 <STO>
A:Cross-references: GB:chr_II; PIDN:AB37940.1; PID:g1707200; GSPDB:GN00020; CESP:C13A10.
C:Genetics:
A:Map position: 3
A:Query Match 6.5%; Score 246.5; DB 2; Length 315;
Best Local Similarity 36.8%; Pred. No. 6.7e-07;
Matches 56; Conservative 27; Mismatches 62; Indels 7; Gaps 3;

QY 475 ESFGLFSCIINGEQTHRAIFRVPVRHEDELEVDPLVQLQAEYDYWEAYNNRTG 534
Db 31 DSSGVSSC-TTSDSQNTHRVQSAFHPRHDPDELLLEIGDAVHVDRTADHWSYCTNLTG 89
QY 535 ARGVFPAYIAIEVTKPE-HMAALAKNSDWV-----DQFRVFKLGSQVQVYHKGNDVLCA 588
Db 90 QSGCIPFASIVCEIDLVEICLGPALPTNATKILSGDRDTEFLTMLASIEVAHHKGNVDLTQ 149
QY 589 AMQKATATTRLTVHFNPPSSCVLEISVRGVKI 620
Db 150 AMNKVLSMKNSBEIIVPQVLMELSPRGHIV 181

RESULT 3
A35363
synapsin I splice form a - human
C:Species: Homo sapiens (man)
C:Date: 17-Aug-1990 #sequence_revision 17-Aug-1990 #text_change 01-Dec-2000
C:Accession: A35363; B35363; A35805
R:Suedhof, T.C.
J. Biol. Chem. 265, 7849-7852, 1990
A:Title: The structure of the human synapsin I gene and protein.
A:Reference number: A35363; MUID:90243651; PMID:2110562
A:Accession: A35363
A:Molecule type: DNA
A:Residues: 1-705 <SUE>
A:Cross-references: GB:M58371; GB:J05431

```

```

A:Accession: B35363
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-659, 'KASPAQAQP' <SU2>
A:Cross-references: GB:M58378; GB:J05431
R:Sauerwald, A.; Hoesche, C.; Oschwald, R.; Killmann, M.W.
J. Biol. Chem. 265, 14932-14937, 1990
A:Title: The 5'-flanking region of the synapsin I gene. A G+C-rich, TATA- and CAAT-le
A:Reference number: A35805; MUID:90368667; PMID:2118519
A:Accession: A35805
A:Molecule type: DNA
A:Residues: 1-125 <SAU>
A:Cross-references: GB:M55301; NID:g338655; PIDN:AAA60608.1; PID:g553654; GB:J05630
C:Genetics:
A:Gene: GDB:SYN1
A:Cross-references: GDB:119606; OMIM:313440
A:Map position: Xp11.23-Xp11.23
C:Keywords: actin binding; alternative splicing; phosphoprotein

Query Match 5.2%; Score 195.5; DB 2; Length 705;
Best Local Similarity 25.8%; Pred. No. 0.0012;
Matches 73; Conservative 25; Mismatches 128; Indels 57; Gaps 10;

QY 100 TGDTPGAEDDEDDDEARARFAGPPKAESGQBPASRGQSQGQSQGSGDYTRPK 159
Db 448 TSQAPGPPAQAQPPPGGGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 506
QY 160 RPTTLNLFPPQVRSQDTLNNSLKKHWSQDRVSRSSSKLTGTEQTPPHEHICLSEELPP 219
Db 507 RLPSPTSAPQASQAAPPTQGG-----RQSRPVAGGPGAPPAARPPASPPQR 556
QY 220 OSGPAPTTDGTSTDSFRRSTATQMAPG-----GPPAAPGGGSHRDRIHYQADVR 274
Db 557 QAGP-POATQTSVSGPAPPKASG--APGGQORQGPQKPPGPPGPPGPPGPPGPPGPP 602
QY 275 LEATEIYLTQVR--PPDAAEPT-----SAPLPPTESRMSVSSDDPAAVPSTAGRRHP 327
Db 603 -QASQ---AGFVRTGPTTQQPRPSGPGAPAKPOLAQPSQDVPPPPATAAAGGPPHP 658
QY 328 SISEEEGFCCLSPERAEPGGGWRGSLCEPPPPPPRASLSSD 370
Db 659 QLNKSQSLTNAFLPE-----PAPRPSPSLSD 685

RESULT 4
B35098
MHC class III histocompatibility antigen HLA-B-associated protein 2 [imported] - huma
C:Species: Homo sapiens (man)
C:Date: 10-Aug-1990 #sequence_revision 06-Nov-1992 #text_change 24-Aug-2001
C:Accession: B35098
R:Banerji, J.; Sands, J.; Strominger, J.L.; Spies, T.
Proc. Natl. Acad. Sci. U.S.A. 87, 2374-2378, 1990
A:Title: A gene pair from the human major histocompatibility complex encodes large pr
A:Reference number: A35098; MUID:90192810; PMID:2156268
A:Accession: B35098
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2142 <BAN>
A:Cross-references: GB:M33509; NID:g179338; PIDN:AAA35585.1; PID:g179339; GB:M31293
A:Note: the authors translated the codon AGT for residue 97 as Gly
C:Superfamily: collagen alpha 1(IV) chain

Query Match 5.1%; Score 194.5; DB 2; Length 2142;
Best Local Similarity 23.2%; Pred. No. 0.0053;
Matches 118; Conservative 33; Mismatches 170; Indels 187; Gaps 22;

QY 13 AASPPAASPFLGLHIASPPNFRUTHDISLEEFED-EDLSEITDECGISLOCKDTLSLRPP 71
Db 1281 APAPRAA-----AKSPDLSNQSQAENEWETASESSDFTSERRGDKERAPPVLLTPK 1334
QY 72 RAGLLSAGGAGSRLOAEMQLMDLIDATGDPGAEDDEEDDEERARRRCPGPPKAES 131
Db 1335 AVGTGGGGGGAAGVGTISA-MSRGDLSQRAKDLKSKRSFSQSRPGRMRQRNRPRGPGKAGSS 1393

```

Db 1454 PSSSAVFLRDQVTHSNPAGTQQALQLSSRQGSVTPAGGHRHKPGPPQAPQG----- 1506

QY 265 DRIHYQADVRLATEEIIYLTQVRPDAAEPTSAFLPTTSMRMSVSSDP----- 313

Db 1507 -----PSRPPTRYEQRV-----NSGLSSDPHFEEPPGPMVRG 1539

QY 314 -----DPAAY-----PSTAGRP--HPSISEEEEGFDCLSSPERAEPPGGWGRSLGEGPPP 361

Db 1540 VGGTPRDSAGVSPFPPKRRRPPRKPELLQEE-----SLP-----PPHSSGFLGSKPEGPG 1590

QY 362 PPRASLSDTSALS---YDSVKYTLVVDEHAQLVELSLRP-----CFGDY 403

Db 1591 PQAESRDTGTATEPHTLWNRL-----HTATSRKSYRPSSEMPHWEPLSPFEDVACTE 1642

QY 404 SDESATSATVYDNCASVSS-----PYESAIGEE 430

Db 1643 MSQSDSGVDLSDGSDQSVSSGPCSQSRSSPDGGLKGAEGPKRPGGSSPLNAVPCEGPPG-- 1700

QY 431 YEEAPRPOPPA 441

Db 1701 -SEPPRRPPPA 1710

RESULT 6

S37671

MHC class III histocompatibility antigen HLA-B-associated protein 2 [similarity] - hu

C:Species: Homo sapiens (man)

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 15-Sep-2000

C:Accession: S37671

R:Bougueleret, L.

submitted to the EMBL Data Library, August 1992

A:Reference number: S37671

A:Accession: S37671

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1870 <BOU>

A:Cross-references: EMBL:z15025; NID:g29374; PID:g29375

C:Genetics:

A:Map position: 6p21.3

A:Introns: 38/2; 97/2; 129/3; 154/1; 202/1; 252/3; 279/2; 327/1; 357/2; 429/3; 588/1;

C:Superfamily: collagen alpha 1(IV) chain

Query Match 5.0%; Score 190.5; DB 2; Length 1870;

Best Local Similarity 23.0%; Pred. No. 0.0076;

Matches 127; Conservative 32; Mismatches 179; Indels 213; Gaps 25;

QY 3 PRES---GCLGGGAASPPAASP-----FLGLHIA-----SPPNRLTHDISLEEF 45

Db 1259 FENAAARGSEKGPLSLTLPASAPGPEALTITVTPAPRRAAKSPDLSNQNSDQANEWE 1318

QY 46 D-EDLSSETTDCGISLOCKDTLSLRPRAGLLSAGGGAGSRLQAEMLQMDLIDATGDT 104

Db 1319 TASESDTTSERGDKEAPPVLLTPKAVCTPGGGGAVPGISA-MSRGDLSQRKDL 1377

QY 105 GAEDDEEDDEERAARRPGAPPKAESQGEPAASRGQSGQSGOGSPGSGDYYRKRPTTL 164

Db 1378 KRFSFSSQRPMEQRNRRPGCGKAGSGS---SSGAGGGPGGRTGPRGD----- 1425

QY 165 NLFPQVPRSQDTLNNSLGKHSQWDRYSRSSPLKTGEOTPPHEHICLSBELP----- 219

Db 1426 -----KRSWSPKRNRRPP-----EERPP-----GLPLPPPPSS 1455

QY 220 -----QSGPAPTTDRGTSTDSPCRRSATQMAPG-----GPPAAPPGGRGSHR 264

Db 1456 SVFRLDQVTHSNPAGIQALQLSS--RQGSVT--APGHPHKKPGPPQAPQG----- 1504

QY 265 DRIHYQADVRLATEEIIYLTQVRPDAAEPTSAFLPTTSMRMSVSSDP----- 313

Db 1505 -----PSRPPTRYEQRV-----NSGLSSDPHFEEPPGPMVRG 1537

QY 314 -----DPAAY-----PSTAGRP--HPSISEEEEGFDCLSSPERAEPPGGWGRSLGEGPPP 361

Db 1538 VGGTPRDSAGVSPFPPKRRRPPRKPELLQEE-----SLP-----PPHSSGFLGSKPEGPG 1590





Qy 604 NPPSSCVLEISVRGKIGVKRADDQSEAKGNKCSHFQ-----LKNISFCGVHPKNNKY--- 656  
Db 1638 -----KVSAGQITL-----TDNQRKLFFRRHYPLNTVTFCDLDPQERKWTKT 1679  
Qy 657 -----FGFIT--KHPADHRFACHVFVSDSTKALAESV 687  
Db 1680 DSGSPAKLFGFVARKQGSTTNDNVCHLFAELDDPQPAIAIV 1719  
RESULT 9  
T26517  
hypothetical protein Y18D10A.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T26517  
R:Harris, B.  
submitted to the EMBL Data Library, December 1998  
A:Reference number: Z20226  
A:Accession: T26517  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1634 <MIL>  
A:Cross-references: EMBL:AL034393; PIDN:CAA22308.1; CESP:Y18D10A.1  
A:Experimental source: clone Y18D10A  
C:Genetics:  
A:Gene: CESP:Y18D10A.1  
A:Introns: 7/3; 44/3; 106/3; 193/1; 500/3; 533/3; 560/3; 709/3; 1218/3; 1318/2; 1  
Query Match 4.8%; Score 182; DB 2; Length 1634;  
Best Local Similarity 21.5%; Pred. No. 0.019;  
Matches 98; Conservative 57; Mismatches 168; Indels 132; Gaps 18;  
Qy 72 RAGLLSAGGGAGS-----RLQAEMLQMDLIDATGDTGCAEDDEDDDD 114  
Db 378 KALLGAAGAAGASEPGSSSIHGKGTRELQTEMKGEV----- 420  
Qy 115 EERARRPGAPPPKAESQEPASRGQSGOGSGDTPYRKPTTLNLPQVPRSQ 174  
Db 421 KEKARM-----RAE-----ASAGRSQAGCPAPAAASELQDP--PQDFGLSMSDFGSD 467  
Qy 175 DTLNNSLGKHKHWDVRSRSPKLTGEQTPPHEHICLSEELPPQSGPAPTDRGTST 234  
Db 468 SETEBHQKQERHI--PAMVTRRSARLSALPVT--KKASSSSKMPPPSPST----- 517  
Qy 235 SPCRES-----TATQWAPGGPPAAGCGRSHRDRIHYQADVLEATEEILYTPVQRP 290  
Db 518 -PGRGRPRRLTSMSEPAATAVTPAPRG---RPRGRSAKVS-ENTE-----PLSEAP 567  
Qy 291 DAAEPTSAFLPPTSRMSVSSDDPDAAYPSTAGRPHPSISEERGFDCI---SSPERAEP 347  
Db 568 SAPYKRGGRPRSRSTMSITEDSEPTSTAAKRSKRAESDEEEQDLKLTNKSPEKPKK 627  
Qy 348 PGGWGRGSLGE-----PPPP-----PPAS 366  
Db 628 PSKTTTEVTGVDLKKRLRDLTKATTATVHTPGPLRTKMRMRAPTAVTSSKKEPKNA 687  
Qy 367 LSSDPSALSYSQVYKTLVVDEHAQELVSLR-----PCRGDYSDESATVYDNCASV 419  
Db 688 GSASSINEEHEDEDTMLEQGT-LDLPQOTSQOEPRISCSGSELLDEQFDA----- 737  
Qy 420 SSPYESAIGEYEAPRPPQACLSSESDTPDEPDV 454  
Db 738 SEHSGTVPSAPELTKNPAPPVPEASEASAEPKPI 772  
RESULT 10  
T03455  
ALR protein - human  
C:Species: Homo sapiens (man)  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 21-Jul-2000  
C:Accession: T03455  
R:Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano,  
Oncogene 15, 549-560, 1997

A:Title: Structure and expression pattern of human ALR, a novel gene with strong homo  
A:Reference number: Z14954; MUID:97388474; PMID:9247308  
A:Accession: T03455  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-4957 <PRA>  
A:Cross-references: EMBL:AF010404; NID:g2358286; PIDN:AAC51735.1; PID:g2358287  
C:Genetics:  
A:Gene: ALR  
A:Map position: 12  
A:Superfamily: human ALR protein  
C:Keywords: alternative splicing  
Query Match 4.7%; Score 179.5; DB 2; Length 4957;  
Best Local Similarity 21.8%; Pred. No. 0.099;  
Matches 110; Conservative 40; Mismatches 166; Indels 189; Gaps 24;  
Qy 1 MAERESGGLGGGAASPPAASPFLGLHIAAPPNFRUTHDISLEEFDEDLSEI-TDCEGIS 59  
Db 1675 LAVPESPGVGGCKASEP-----LLSPPPFG-----ESRKALEVKKEELGAS 1715  
Qy 60 LOCKDTLSLRPPRAGLLSAGGG---AGSRLOAEMLQMDLIDATGDTGCAEDDEDDDEE 116  
Db 1716 -----SPSYGPPNMGFVDSPPSGTHLGGLEK-----TP-----DVF 1747  
Qy 117 RAARRPGAPPPKAESQEPASRGQSGOGSGDTPYRKPTTLN-LFPQVPRSOD 175  
Db 1748 KAPLTPRA-----SQVEPQSPGLG--LRQEPPPAQAALAPSPSPHPD 1787  
Qy 176 TLNNSLGKHKHWDVRSRSPKLTGEQTPPHEHICLSEELPPQSGP----- 223  
Db 1788 IFRPGSYTDPY-----AQPLTPRPQPPPESSC---ALPPRSLSDPFSRVPVSPQ 1836  
Qy 224 -----APTDTGRTSDSPCRSTATQMAP--CGPPAAPGCGRSHRDRIHQADV 274  
Db 1837 SOSSSOSPLTPRPLSAEAFPCSPVTPRFQSPDPYSRPPSRP-----QSRDP----- 1882  
Qy 275 LEATEEILYTPVORPPDAAEPTSAFLPPTSRMSVSSDDPDAAYPS-TAGRPHPSISEE 333  
Db 1883 -----FAPLHKPPRPOPPEVAFKAGSLAHTSLGAGGFPAALPAGPAGELHAKVP 1934  
Qy 334 EGPDCUSSPERAEPGGG-----WRGSLGEP---PPPPRASLSSDTSALSY--- 376  
Db 1935 -----PNFVRSPTGAFVGTPTSPMRFTFPQAVGEPSLKPPVPQGLPPPHGINS 1987  
Qy 377 -----DSVKYTLVVDEHAQELVSLRPCFGDYSDESATVYDNCASVSPYESAIG 428  
Db 1988 PGTLGKPGQSTNTTVATGN-----FHPSGSPGLPSSGGTGC 2022  
Qy 429 EYEEAPRPPQACLSSESDTPDEPD 453  
Db 2023 ESYGLSPL-RPPSVL---PPAPD 2042  
RESULT 11  
T03454  
ALR protein - human  
C:Species: Homo sapiens (man)  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 21-Jul-2000  
C:Accession: T03454  
R:Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Ya  
Oncogene 15, 549-560, 1997  
A:Title: Structure and expression pattern of human ALR, a novel gene with strong homo  
A:Reference number: Z14954; MUID:97388474; PMID:9247308  
A:Accession: T03454  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-5262 <PRA>  
A:Cross-references: EMBL:AF010403; NID:g2358284; PIDN:AAC51734.1; PID:g2358285  
C:Genetics:  
A:Gene: ALR  
A:Map position: 12  
C:Superfamily: human ALR protein

C:Keywords: alternative splicing

Query Match 4.7%; Score 179.5; DB 2; Length 5262;  
Best Local Similarity 21.8%; Pred. No. 0.11;  
Matches 110; Conservative 40; Mismatches 166; Indels 189; Gaps 24;  
QY 1 MAERESGGLGGGAAPPAAPFLGLHIAAPPNFRLTHDISLEEFDELSRI-TDECGIS 59  
Db 1980 LAVPESGVGGKASRP-----LLSPDPFG-----ESRKALEVKKEELGAS 2020  
QY 60 LOCKOTLSLRPRAGLLSAGGG---AGSRLOAEMQLMDLIDATGDTGAEDDEDDDEE 116  
Db 2021 -----SPSYGPNLGFVDPSSGTHLGLLEK-----TP-----DVF 2052  
QY 117 RAARPGAGPPRAESQEPASRGOGSOGSOGCGSDTYRKPRTTILN-LFPQVPRSQD 175  
Db 2053 KAPLTPRA-----SQVEPQSLG--LRQEPPPAALAPSPSPHD 2092  
QY 176 TLNNSLGGKHSWQDRVSRSSSLKTGEQTPPHEHICLSEELPQSGP----- 223  
Db 2093 IFRPGSYTDPY-----AQPLTPRPQPPPEPSCC---ALPRLSPDFSRVPVSPQ 2141  
QY 224 -----APTDRGTSTDSPCRSTATQAPP---CGPPAAPGGRGHSHRDRIHYQADV 274  
Db 2142 SOSSSOPLTRPLSAEAFPCSPVTRFQSDPDVYSRPSRP-----QSRDP----- 2187  
QY 275 LEATEEIVLPVQRPDAAEPTSAFLPPTESRMSVSSDDPAAYPS-TAGRPHPFSIEEE 333  
Db 2188 -----FAPLHKPRQPPEVAFAKAGSLAHTSLGAGGFPAALPAGAGELHAKVPSGQ 2239  
QY 334 EGFDCLSPERAEPGGG-----WRGSLGEP---PPPRASLSDTSALSYS-- 376  
Db 2240 -----PPNFVRSPTGAFVGTPTSPMRFTFPQAVGEPSLKPPVPQGLPPHGINSHFG 2292  
QY 377 -----DSVKYTLVVDEHAQLELVSLRCPGCDYSDSDSATVYDNCASVSSYSAIG 428  
Db 2293 PGFTLGKQSTNYTVATGN-----FHPGSGPLGSPSGSGTG 2327  
QY 429 EYEEAPRPQPPACLSDESDTDEPD 453  
Db 2328 ESYGLSPL-RPPSVL-----PPAPD 2347  
RESULT 12  
A54970  
tensin, cardiac muscle - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 21-Jul-2000  
C:Accession: A54970; S38330; S21544  
R:Lo, S.H.; An, Q.; Bao, S.; Wong, W.K.; Liu, Y.; Janney, P.A.; Hartwig, J.H.; Chen, L.H.  
J. Biol. Chem. 269, 22310-22319, 1994  
A:Title: Molecular cloning of chick cardiac muscle tensin. Full-length cDNA sequence, ex  
A:Reference number: A54970; MUID:94350987; PMID:8071358  
A:Accession: A54970  
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
A:Molecule type: mRNA  
A:Residues: 1-1744 <LOA>  
A:Cross-references: GB:M96625  
R:van de Werken, R.; Gennari, M.; Tavella, S.; Bet, P.; Molina, F.; Lin, S.; Cancedda, P  
Eur. J. Biochem. 217, 781-790, 1993  
A:Title: Modulation of tensin and vimentin expression in chick embryo developing cartila  
A:Reference number: S38330; MUID:94039118; PMID:8223621  
A:Accession: S38330  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1469-1744 <VAN>  
A:Cross-references: EMBL:X66286; NID:g63802; PIDN:CAA46992.1; PID:g63803  
C:Superfamily: SH2 homology  
C:Keywords: cardiac muscle; heart  
F:1472-1581/Domain: SH2 homology <SH2>

Query Match 4.7%; Score 179; DB 2; Length 1744;  
Best Local Similarity 19.8%; Pred. No. 0.031;

Matches 177; Conservative 97; Mismatches 312; Indels 306; Gaps 39;

QY 13 AASPPAASFLGLHIAAPPN-----FRLTHDISLEEFDELSRI-TDECGISLQCKDTL 66  
Db 928 ASYSPAGSQ--QLLVSPSPPTAPASQPLPK-GLSEYED-----LSRSGEPL 973  
QY 67 SLRPPRAGLLSAGGGAGSLOAEMQLMDLIDATGDTGAEDDEDDDEERAARRP---- 122  
Db 974 NLE-----GLVAHRVAGVQSREKSP--EESTVPARRTPSDSHYEKSPSPGSPRSTVL 1027  
QY 123 -----GAGPPK-----AESGOEPASRGOGSOGSOGCGSDT-----YR 157  
Db 1028 PEVVSTIAANPGGRPKPEHLHSYKEAFEMESASPSLSTGGVRSPPGLAKTFLSALGLK 1087  
QY 158 PKRPTTILNLF-----QVPRS-----QDTLNN 179  
Db 1088 PHPADILLHPVGELECEAGADSEEPRSYVESVARTATTGRAGNLPAAPQVGVLEYPARN 1147  
QY 180 NSLGKHKHSWQDRVSRSSSLKT-----GEOTPPHEHICLSEELPQSGP 223  
Db 1148 GATGNSFTVPSPVS--TSSPHSVDSGLRSYSESGPHGTIVTPPHA-VAETAYRSPWVSQ 1205  
QY 224 APTDRGTSTDSP-----PQGP-----PAAPPGRGHSHRDRIHY 269  
Db 1206 TPSAHSSTQTSPPSFQAGTLGSPYASPDYPDGRGGFQDPQARQOPQVSVVGHALPGS 1265  
QY 237 ---CRSTATQMAP-----PQGP-----PAAPPGRGHSHRDRIHY 269  
Db 1266 PRTLHRTVATNTPPSGFGRRAANPAVASVPGSPGLGRHTVSPHAPPGSPSLAR-----HQ 1321  
QY 270 QADVRLAEIEIYLTVPQRPDAAEPTSA--FLPPTESRMSV-----SSDDPP 315  
Db 1322 MAAPVPGSPMYGSSPEERRPTLSRQSSASGYQPPSTPFPVSPYAYPGTSPHSSPDS 1381  
QY 316 AAYPTAGRPHPSISEEEGFCCLSPERAEPGGGWRGSL-----GEPPPPPRASLS 368  
Db 1382 AAYRQGSPTQPALPEKRR---MSAGERS-----NSLPNYATVNGKASSPLSSGMS 1429  
QY 369 SDTSALSVDYKVTLLVVDEHAQLELVSLRCPGCDYSDSDSATVYDNCASVSSP---YE 424  
Db 1430 SPSSG-SAVAFSHTL-----PDFSKSPMDISPETRANKVQVQDTSKYWTK 1474  
QY 425 SAIGEEYEEAPRPQPPACLSDESDTDEPDVHFSSKFLNVMFMSGRSSRSASAEGLFSCII 484  
Db 1475 PDLRQAIA-----LLKOREPGAFIIRDSHSPRGAYGLAMKVASPPP-----TVMQ 1521  
QY 485 NGEQEQTHRALFRV-----PRHEDELEVEDDPLLVLEQAEQYWEAYNMTGARGVFP 540  
Db 1522 QNKKGDITNELVRHFLIETSPRGV-KLKGCPNEPNFGCLSLALVYQHSIMPLALPCKLVIP 1580  
QY 541 AYYAIEVTEPEHMAALAKNS--DWVDQ---FRVKFLGSOVQPYHKGNDVLCANMQKI-- 593  
Db 1581 DRDPMEEKD---AASTTNSATDLLKQGAACNVLFINSVEMESITGPOAISKAETLV 1636  
QY 594 --ATTRRLTVHFNPPSSCVLEISVRGKIGVADDSQEAQKNGKSHFFQ-----LKNISFC 647  
Db 1637 ADPTPTATIVHF-----KVSAGQITL-----TDNQKLFERHVPPLNTVTF 1678  
QY 648 GYHPKNKY-----FGFIT-KHPADHRFACHVFSVSDSTKALAESV 687  
Db 1679 DLDPQERKWKTKDGGSPAKLFGFVARKQGSTTDNVCHLFAELDPDQAAAAIV 1730

RESULT 13  
A57075  
tensin - chicken (fragment)  
C:Species: Gallus gallus (chicken)  
C:Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 21-Jul-2000  
C:Accession: A57075  
R:Chuang, J.-Z.; Lin, D.C.; Lin, S.  
J. Cell Biol. 128, 1095-1109, 1995  
A:Title: Molecular cloning, expression, and mapping of the high affinity actin-cappin  
A:Reference number: A57075; MUID:95204530; PMID:7896874

A:Accession: A57075

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-1792 &lt;CHU&gt;

A:Cross-references: GB:L06662; NID:g212754; PIDN:AAA73949.1; PID:g212755

C:Superfamily: SH2 homology

F:1520-1629/Domain: SH2 homology &lt;SH2&gt;

Query Match 4.7%; Score 179; DB 2; Length 1792;  
Best Local Similarity 19.3%; Pred. No. 0.032;  
Matches 171; Conservative 108; Mismatches 311; Indels 296; Gaps 38;

QY 13 AASPAPASPELGLHTASPPN-----ERLHDISLEFEDEDLSEITDEGISLQCKDTL 66  
DB 976 ASYPAGSQ--QLLVSSPPSTAPAQOLPHK-CLESYD-----LSRSEEL 1021  
QY 67 SURPPRAGLSAGGGAGSRQAEMQLMDLIDATGDTFGAEDDEDDERAAARRP----- 122  
DB 1022 NLE-----GLVAHRVAGVQREKSP--EESTVPARRRTPSDSHVEKSPSPRSTVLS 1075  
QY 123 -----GAGPPK-----AESQGPASRGQSQGQSQGPGSDT-----YR 157  
DB 1076 PEVVTIANPGRRKPEHLHSYKEAFEMESASPSLTSGGVKSPPLAKTFLSALGLK 1135  
QY 158 PKRPRTLLMFP-----QVPRSS-----QDTLNN 179  
DB 1136 PHNPADILLHPVELEGAGADEEPRSYVESVARTATTGRAGNLPAAGQVGLVPAARN 1195  
QY 180 NSLGGKHSQDRVSRSSPLKT-----GEOTPPHEHICLSEELPPQSGP 223  
DB 1196 GAFGNSFTVPSPVS--TSSPIHSVDGASLRSPSGPHGTVTTPHA-VAETAYRSPMVQ 1253  
QY 224 APTDRTGTSTDP-----PGGP-----SSDPP 315  
DB 1254 TFSANSSVTQTSPPSFQAGTLGSPYASPDYDGRAGTQPDQAQOQVSVGVHLPQS 1313  
QY 237 -----CRSTATOMAP-----PAAPGGRGHSHRDIHY 269  
DB 1314 PRTLHRTVATNTPSPGFRRAANPAVASVPGSLGRHTVSPHAPGPSLAR-----HQ 1369  
QY 270 QADVLEATEEILYTPVQPDAAEPTSA--FLPPTESRMV-----SSDPP 315  
DB 1370 MAAPVPGPMYGYSPERKRTLRQSSASGYQPPSPFPVPAFYPGTSTHSSSPDS 1429  
QY 316 AAYPTAGRPSPISIEEEDGDCLSSEPERAEPGGWRSGL-----GPPPPPRASLS 368  
DB 1430 AAYRGSPTQPALPEKRR-----MSAGERS-----NSLPNATVNGKASSPLSGMS 1477  
QY 369 SDTSALSDVSKYTLVDEHAQLLELVLRCFGDYSDSDSATVYDNCASVSPYSAG 428  
DB 1478 SPSSG---SAVAFSHTLPDFSKFMPDISP-----ETRANVKEVDTSKY 1519  
QY 429 EYEEAPRPPACLSDESTDPEDVHFSSKFLNVMSGRSSSAESFGLFSCIINGEE 488  
DB 1520 WTKPDISRQOAILL-KDREGCAFIIRDSHSFRGAYGLAMKVASPPP-----TVMOQNK 1573  
QY 489 QBOHRAIRFV-----PRHEDELEVDPLLLVELQAEIYWEYAYNNMRTGARGVFAYXA 544  
DB 1574 GDITNELVRHFLIETSPRGV-KLKGCPNEPFGCLSAVYQHSIMPLALPKLVIPODRP 1632  
QY 545 IEVTEPEHMAALAKNS--DWVDQ---FRVKFLGSVQVPHKGNVDLCAMQKI-----AT 595  
DB 1633 MEKKD-----AASTNSATDILLKOGAACNLVFINSVEMESLTGPQAIKSAVAETLVADPT 1688  
QY 596 TRLRVHFNPPSSCVLEISVRGVIGKADDSQEAQKNCSHFFO-----LKNISFCGYHP 651  
DB 1689 PRATIVHF-----KVSAGIYL-----TDNQKLFRRHYPLNTVTFCDLDP 1730  
QY 652 KNNKY-----FGFIT-KIPADHRFACHVFVSEDTKALAESV 687  
DB 1731 QERKWTKTGSGPAKLFQGVARKQGGTTEENVCHFLFAELDPQAAAAIV 1778

RESULT 15

RESULT 14

JC5209

Insulin receptor substrate 1 - chicken

C:Species: Gallus gallus (chicken)

C:Date: 20-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 16-Jul-1999

C:Accession: JC5209; PC4305

R:Taouis, M.; Taylor, S.I.; Reitman, M.

Gene 178, 51-55, 1996

A:Title: Cloning of the chicken insulin receptor substrate 1 gene.

A:Reference number: JC5209; MUID:97080546; PMID:8921891

A:Accession: JC5209

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-1240 &lt;TAO&gt;

A:Cross-references: GB:043502; NID:g1685084; PIDN:AA60050.1; PID:g1685085

A:Accession: PC4305

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-5 &lt;TA2&gt;

C:Comment: This protein acts as a docking protein and mediates multiple interactions

C:Genetics:

A:Gene: ist-1

C:Superfamily: insulin receptor substrate IRS-1; pleckstrin repeat homology

C:Keywords: phosphoprotein

F:11-112/Domain: pleckstrin repeat homology &lt;PLK&gt;

F:463,549,610,630,660,730,940,987,1010/Binding site: phosphate (Tyr) (covalent) #stat

Query Match 4.7%; Score 177.5; DB 2; Length 1240;

Best Local Similarity 22.3%; Pred. No. 0.025;

Matches 129; Conservative 59; Mismatches 219; Indels 171; Gaps 26;

QY 2 AERESGGGG-----GAASPP-----AASP-----FLGLHIASPPN 32  
DB 308 ATSPAGGVGGFPSSFRVRESSDGEETMSPPDQDQSPVSPSTNRTAHRHRSALLQPP- 366  
QY 33 FRLHDIHL-----EEFEDELSE-----ITDEGILSQCKDTLSLRPPRAGLL 76  
DB 367 --LNHSRIFMPASRCSPSATSPVLSLSSSTSGHGSTSDCLFPRRSASVSGSPSGGFI 424  
QY 77 SAGGGAGSRQAEMQLMDLIDATGDTFGAEDDEDDERAAARRPCAGPP--KAESQGE 134  
DB 425 SDEYSSPCDFRSFRSVTPDLSLHTPPARGEEE---LSNYICMGKGKGPSTLTAPNGHY 481  
QY 135 PASRGQSQ 190  
DB 482 ILSRG---ANGHRCPTGT-----LGTSPALAGDEQTSAAADLDNFRKRTH--- 524  
QY 191 RVSRSSPLKTEQTPPHEHICLSEELPPQSGGPAPTTDGTSTDSPCRRSTATQMAPPG 250  
DB 525 --SAGTSTIITHTQPSQSSVASEEY-----TEMM----- 553  
QY 251 PPAAPPGG--RGHSHRDKRIHYQADVRLAETEEIYLTVPQRPDAAEPTSAFLPPTESRMS 308  
DB 554 -PAYPPGGSGGRAMPGLSAFVTRSYRPEGLMHPLERRAGTSPDSTLHTDDGYMP 612  
QY 309 VSSDDPDAAPYPTAGR-----PHPSIEEPEGDFCLSSP-ERAEPGP-----GG 351  
DB 613 MS-----PGWPPVSSGRKNGDYMSPKSVSTPQOITNSMTDPSERVYPNGYMMHSRG 668  
QY 352 WRGSLGEPDPPPRASLSDDTSALSYDSVKYTLVYDEHAQL-----ELVSLRPECFG 401  
DB 669 CSPDGGGSSSSSSSNAPVSGTSGKLTWNGVGGHSHVLPHPKPPVSSGSGKLLPCTG 728  
QY 402 DYSD---ESDSATVYDNCASVSPYSIAIGEEYEAAPQP-----PACLSSEDTTPE 451  
DB 729 DYMDYVPGDSNT-----SDPSLSYG--PEDQHKPVLVSLYSLPDAFKHPSPGE 777  
QY 452 PDVHFSKFLNV-FMSGR-----SRSSSAESFG 478  
DB 778 PEEGARHOLRLSTSSGRLLYAATADSSSSSTSSDTLG 815

JS0670

insulin receptor substrate-1 - human

C:Species: Homo sapiens (man)

C:Date: 30-Jun-1992 #sequence\_revision 23-Aug-1996 #text\_change 16-Jul-1999

C:Accession: F53160; JS0670; P0678

R:Araki, E.; Sun, X.J.; Haeg, B.L.; Chuang, L.; Zhang, Y.; Yang-Feng, T.L.; White, M.;

Diabetes 42, 1041-1054, 1993

A:Title: Human skeletal muscle insulin receptor substrate-1. Characterization of the cDN

A:Reference number: F53160; MUID:93292738; PMID:8513971

A:Accession: F53160

A:Status: translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-1242 &lt;RES&gt;

A:Cross-references: GB:562539; NID:9386256; PIDN:AAB71175.1; PID:g386257

R:Nishiyama, M.; Wands, J.R.

Biochem. Biophys. Res. Commun. 183, 280-285, 1992

A:Title: Cloning and increased expression of an insulin receptor substrate-1-like gene i

A:Reference number: JS0670; MUID:92181456; PMID:1311924

A:Accession: JS0670

A:Molecule type: mRNA

A:Residues: 1-134, G', 135-361, R', 363-383, R', 385-1242 &lt;NIS&gt;

A:Cross-references: GB:585963; NID:9246465; PIDN:AAB21608.1; PID:g246466

A:Experimental source: hepatocellular carcinoma cell line FOCUS

R:Smith, L.K.; Bradshaw, M.; Croall, D.E.; Garner, C.W.

Biochem. Biophys. Res. Commun. 196, 767-772, 1993

A:Title: The insulin receptor substrate (IRS-1) is a PEST protein that is susceptible to

A:Reference number: P0678; MUID:94059102; PMID:8240352

A:Accession: P0678

A:Molecule type: protein

A:Residues: 1-14; 338-353; 411-439; 538-545, 'V', 547-567; 656-697; 724-758; 932-943; 1028-1056; 1

C:Comment: This protein appears to be the major substrate for insulin-stimulated tyrosin

o dock various proteins containing the phosphotyrosine-binding Src-homology domain 2 (SH

C:Comment: This protein contains at least 11 PEST regions, which suggests rapid turnover

C:Genetics:

A:Gene: GDB: IRS1

A:Cross-references: GDB:133974; OMIM:147545

A:Map position: 2q36-2q36

A:Introns: #status absent

C:Superfamily: insulin receptor substrate IRS-1; pleckstrin repeat homology

C:Keywords: ATP; phosphoprotein

F:11-113/Domain: pleckstrin repeat homology &lt;PLK&gt;

F:46,465,551,612,632,662,732,941,989,1012/Binding site: phosphate (Tyr) (covalent) #stat

F:78,527,1100,1223/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #

F:161/Binding site: ATP (Lys) #status predicted

F:189,323,441,624,636,795,920,984,1084,1218/Binding site: phosphate (Ser) (covalent) (by

F:300,351,774/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status pre

Query Match

Best Local Similarity 4.7%; Score 177.5; DB 2; Length 1242;

Matches 117; Conservative 57; Mismatches 213; Indels 123; Gaps 22;

QY 53 TDECGISLQKDTLSLRPRAGLLSAGGGAGSRLOQLMDLIDATGDTPGAEDDED 112

DB 403 TSDCLFPRRSASVSGSDGFTSSDEYSSPCDFRSFRSVTPDSLGLHTPPARGEED- 461

QY 113 DDEERARRPGAGPP--KAESGQEPASRGQSQGQSGQSGDGYRKRPTTLNLFQV 170

DB 462 --LSNVICMGKGSTLTAPNGHYILSRG--GNHRCITPGTG-----LGTSPALAGDE 510

QY 171 PRSQTLNNNSLGKSHQDRVSRSSPLKTEQTPPHEHICLSELPQSGPAPTDRG 230

DB 511 AASADLDNFRFKRTH-----SAGTPTTHQTPSQSSVASIEY----- 551

QY 231 TSTDSPCRRTATQMAPGGPAPPGGR--GHSRDRIRHYQADVLEATEEIVLTPVQR 288

DB 552 -----TEMM-PAYPPGGSGGRLPGHRSFAV---PTRSYPEEGLEHPLER 594

QY 289 PDAAEPTSAFLPPTESRMSVSDPDPAAYPSTAGR-----PHPSISEEEGFDCL 339

DB 595 RGGHRRPDSSTLHTDDGYMPS-----FCVAVPSPGRSGDGYMPSKVSAPQIINPI 650

QY 340 -SSPERAPF-----PGGWGRGSLGEPppppppRASLSSDTALSVDYKVTLLVVDHAOL 391

DB 651 RRHFQVRDPNGYMMSPSGGCGSPDIGGPPSSSSSSNAVPSTSYGKLWTVGGHHSHV 710

QY 392 -----ELVSLRPFCDGYSDES---DSATVVDNCASVSPVESAIGEEYEAPRPQ 438

DB 711 LPHPKPPVESGGGKLLPCTGDMNNSPVGDSNT-----SSPDCYVG---PEDPQHK 759

QY 439 P-----PACLSESDSTPDDEPDVHFKKFLNV-FMSGR-----SRSSAESFGLF 480

DB 760 PVLISYSLPRSEKHTQRPGEPEGARHQHRLSTSGRLLYAATADSSSTSSDSLGGG 819

QY 481 SCIIINGEEQET-----HRAIFRFPVRHED 505

DB 820 YC---GARLEPSLPHPHQVQLPHLPKVD 846

Search completed: December 24, 2002, 08:51:11

Job time : 47 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2002, 08:48:45 ; Search time 36 Seconds  
(without alignments)  
819.157 Million cell updates/sec

Title: US-09-966-561-2  
Perfect score: 3792  
Sequence: 1 MAERESGLGGGAASPAAS.....QQFYKQFVEYTCPTEDIYLE 711

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3789	99.9	711	JIPL_HUMAN	Q9w4f2 homo sapien
2	3549	93.6	707	JIPL_MOUSE	Q9wv19 mus musculu
3	3544.5	93.5	708	JIPL_RAT	Q9r237 r c-jun-ami
4	1292.5	34.1	830	JIP2_MOUSE	Q9ere9 mus musculu
5	1280.5	33.8	824	JIP2_HUMAN	Q13387 homo sapien
6	537.5	14.2	490	SP51_DROME	Q9w0k0 drosophila
7	195.5	5.2	705	SYN1_HUMAN	P17600 homo sapien
8	194.5	5.1	2142	BAT2_HUMAN	P48634 homo sapien
9	193.5	5.1	503	WAIP_HUMAN	O43516 homo sapien
10	184.5	4.9	706	SYN1_BOVIN	P17599 bos taurus
11	179	4.7	1744	TENS_CHICK	O04205 gallus gall
12	177.5	4.7	1242	IRSL_HUMAN	P35568 homo sapien
13	174.5	4.6	1077	HUES_DROME	Q02308 drosophila
14	170.5	4.5	415	SYN1_CANFA	O62732 canis fami
15	170.5	4.5	1664	SLP1_CLOTM	O62852 clostridium
16	170	4.5	3149	TEGU_EBV	P03186 epstein-bar
17	169.5	4.5	1185	DRPL_HUMAN	P54259 homo sapien
18	169	4.5	671	VINE_HUMAN	O60504 homo sapien
19	169	4.5	1233	IRSL_MOUSE	P35569 mus musculu
20	167.5	4.4	2167	SHK1_RAT	Q9wv48 rattus norv
21	166.5	4.4	1235	IRSL_RAT	P35570 rattus norv
22	166	4.4	704	SYN1_RAT	P09951 rattus norv
23	165.5	4.4	548	ERF_HUMAN	P50548 homo sapien
24	165	4.4	1083	T2D3_HUMAN	O00268 homo sapien
25	165	4.4	1291	PERL_MOUSE	O35973 mus musculu
26	165	4.4	1461	IE18_PRIVIF	P11675 pseudorabie
27	163.5	4.3	551	ERF_MOUSE	P70459 mus musculu
28	163.5	4.3	839	APB1_RAT	Q35430 rattus norv
29	162	4.3	1183	DRPL_RAT	P54258 rattus norv
30	159.5	4.2	707	DREB_RAT	Q07266 rattus norv
31	159.5	4.2	1324	IRS2_HUMAN	Q9y4h2 homo sapien
32	159	4.2	2424	CCAA_RABIT	P27884 oryctolagus
33	158.5	4.2	1446	IE18_PRIVKA	P33479 pseudorabie

34	157.5	4.2	561	1	3BP2_HUMAN	P78314 homo sapien
35	157.5	4.2	1125	1	S24C_HUMAN	P53992 homo sapien
36	157	4.1	234	1	PRPM_HUMAN	P10161 homo sapien
37	156.5	4.1	837	1	AFB1_HUMAN	Q02410 homo sapien
38	156.5	4.1	1790	1	SEPA_EMENI	P78621 emericeia
39	156	4.1	670	1	SYN1_MOUSE	O88935 mus musculu
40	155.5	4.1	924	1	104K_THEPA	P15711 theileria p
41	154.5	4.1	568	1	RGS3_MOUSE	Q9dc04 mus musculu
42	154.5	4.1	1174	1	KPCL_COCHE	O42632 cochliobolu
43	154.5	4.1	1443	1	SVJ2_HUMAN	O15056 homo sapien
44	154	4.1	733	1	VINE_MOUSE	Q9r128 mus musculu
45	154	4.1	777	1	RTN1_RAT	Q64548 rattus norv

ALIGNMENTS

RESULT 1  
JIPL\_HUMAN  
ID JIPL\_HUMAN STANDARD; PRT; 711 AA.  
AC Q9UQF2: O43407;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE C-jun-amino-terminal kinase interacting protein 1 (JNK-interacting  
DE protein 1) (JIP-1) (JNK MAP kinase scaffold protein 1) (Islet-brain-1)  
DE (IB-1) (Mitogen-activated protein kinase 8-interacting protein 1).  
GN MAPK8IPI OR JIP1 OR IB1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Insulinoma;  
RX MEDLINE=99134298; PubMed=9933567;  
RA Mooser V., Maillard A., Bonny C., Steinmann M., Shaw P., Yarnall D.P.,  
RA Burns D.K., Schorderet D.F., Nicod P., Waerber G.;  
RT "Genomic organization, fine-mapping, and expression of the human  
RT islet-brain 1 (IB1)/C-jun-amino-terminal kinase interacting protein-1  
RT (JIP-1) gene.";  
RT Genomics 55:202-208(1999).  
[2]  
SEQUENCE OF 468-711 FROM N.A.  
RC TISSUE=Brain;  
RX Yu W., Sarginson J., Gibbs R.A.;  
RT Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
[3]  
RHOGEF BINDING, AND SUBCELLULAR LOCATION.  
RX MEDLINE=20044776; PubMed=10574993;  
RA Meyer D., Liu A., Margolis B.;  
RT "Interaction of c-Jun amino-terminal kinase interacting protein-1 with  
RT p190 rhoGEF and its localization in differentiated neurons.";  
RL J. Biol. Chem. 274:35113-35118(1999).  
[4]  
MUTAGENESIS, AND KINESIN BINDING.  
RX MEDLINE=21135887; PubMed=11238452;  
RA Verhey K.J., Meyer D., Deehan R., Blenis J., Schnapp B.J.,  
RA Rapoport T.A., Margolis B.;  
RT "Cargo of kinesin identified as JIP scaffolding proteins and  
RT associated signaling molecules.";  
RL J. Cell Biol. 152:950-970(2001).  
[5]  
VARIANT NIDDM ASN-59.  
RX MEDLINE=20164330; PubMed=10700186;  
RA Waerber G., Delplanque J., Bonny C., Mooser V., Steinmann M.,  
RA Widmann C., Maillard A., Miklosy J., Dina C., Hani E.H., Vionnet N.,  
RA Nicod P., Boutin P., Froquel P.;  
RT "The gene, MAPK8IPI, encoding islet-brain-1, is a candidate for type 2  
RT diabetes.";  
RL Nat. Genet. 24:291-295(2000).  
CC -!- FUNCTION: The JNK-interacting protein (JIP) group of scaffold  
CC proteins selectively mediates JNK signaling by aggregating

specific components of the MAPK cascade to form a functional JNK signaling module. JIP1 is required for JNK activation in response to excitotoxic stress (By similarity). Cytoplasmic JIP1 causes inhibition of JNK-regulated activity by retaining JNK in the cytoplasm and by inhibiting the JNK phosphorylation of c-Jun (By similarity). May also participate in ApoER2-specific reelin signaling. Directly, or indirectly, regulates GLUT2 gene expression and beta-cell function (By similarity). Appears to have a role in cell signaling in mature and developing nerve terminals.

-! SUBUNIT: Forms homo- or heterooligomeric complexes. Binds specific components of the JNK signaling pathway namely JNK, MAPK7 and MLK2, MLK3 and DLK (By similarity). Also binds the proline-rich domain-containing splice variant of apolipoprotein E receptor 2 (ApoER2) (By similarity). Binds the TPR motif-containing C-terminal of kinesin light chain. Interacts, via the PID domain, with rhoGEF. Binds the cytoplasmic tails of LRP1 and LRP2 (Megalyn) (By similarity).

-! SUBCELLULAR LOCATION: Cytoplasmic. Accumulates in cell surface projections (By similarity). Under certain stress conditions, translocates to the perinuclear region of neurons (By similarity). In insulin-secreting cells, detected in both the cytoplasm and nucleus (By similarity).

-! TISSUE SPECIFICITY: Highly expressed in brain. Expressed in neurons, localizing to neurite tips in differentiating cells. Also expressed in the pancreas, testis and prostate. Low levels in heart, ovary and small intestine.

-! PFM: Phosphorylated by JNK in differentiated cells.

-! DISEASE: Defects in MAPK8IP1 may contribute to autosomal dominant non-insulin-dependent diabetes mellitus type II (NIDDM), which is characterized by an autosomal dominant mode of inheritance, onset during childhood (usually before 25 years of age) and a primary defect in insulin secretion.

-! SIMILARITY: BELONGS TO THE JIP SCAFFOLD PROTEINS FAMILY.

-! SIMILARITY: CONTAINS 1 SH3 DOMAIN.

-! SIMILARITY: CONTAINS 1 PID DOMAIN.

-----

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; AF074091; AAC20443.1; -;  
EMBL; AF071334; AAC19150.1; -;  
TRANSFAC; T04465; -;  
Genew; HGNC:6882; MAPK8IP1.  
MIM; 604641; -;  
MIM; 125853; -;  
InterPro; IPR0000050; PID\_domain.  
InterPro; IPR001452; SH3.  
Pfam; PF00018; SH3; 2.  
Pfam; PF00640; PID; 2.  
ProDom; PD0000066; SH3; 1.  
SMART; SM00462; PTB; 1.  
SMART; SM00326; SH3; 1.  
PROSITE; PS01179; PID; 1.  
PROSITE; PS50002; SH3; 1.  
SH3 domain; Phosphorylation; Disease mutation; Diabetes mellitus.  
KW DOMAIN 42 48  
FT DOMAIN 79 84  
FT DOMAIN 107 116  
FT DOMAIN 127 285  
FT DOMAIN 331 334  
FT DOMAIN 359 363  
FT DOMAIN 488 549  
FT DOMAIN 561 700  
FT VARIANT 59 59  
FT MUTAGEN 704 704  
FT MUTAGEN 709 709  
FT SEQUENCE 711 AA; 777524 MW; 55EA53B30080A751 CRG64;  
SQ







FT VARSPLIC 1 33 MAERESGLSGGAASPPAASPFLGLHIASPPNFR -> MQLV  
 FT LKMSPPDNDMSLEDDQWERW (IN ISOFORM 2).  
 FT MISSING (IN ISOFORM 3).  
 FT E -> EPMQVQLQVLDLEIKRAAEQKLISEEDLNAA  
 FT (IN ISOFORM 3).  
 FT CONFLICT 38 I -> V (IN REF. 2).  
 FT CONFLICT 149 S -> C (IN REF. 1).  
 SQ SEQUENCE 708 AA; 77317 MW; 4923FD55F1C511F4 CRC64;

Query Match 93.5%; Score 3544.5; DB 1; Length 708;  
 Best Local Similarity 94.0%; Pred. No. 7.5e-172;  
 Matches 669; Conservative 11; Mismatches 27; Indels 5; Gaps 4;

QY 1 MAERESGLSGGAASPPAASPFLGLHIASPPNFRFLTHDLSLEFEDEDELSITDEGIGSL 60  
 DB 1 MAERES-GLSGGAASPPAASPFLGLHIASPPNFRFLTHDLSLEFEDEDELSITDEGIGSL 59  
 QY 61 OCKDTLSLRPRAGLLSAG-GGAGSRLQAEMLQMDLIDATGTPCAEDEDDEDERAA 119  
 DB 60 OCKDTLSLRPRAGLLSAGSGSRLQAEMLQMDLIDASTPQAEDEDEDD-ELAA 118  
 QY 120 RRGAGPPKAESGPPASRGQSGQSGGSDTYRKRPTTLNLFQVPRSQDTLNN 179  
 DB 119 ORPGVGSKAESGPPASRSQG--QGQPGTSGDTPYRKRPTTLNLFQVPRSQDTLNN 176  
 QY 180 NSLKKHSHQDRVSRSSPLKGTGTPPHEHICLSLPELPGQSPVPTQGRSTSDSPCR 239  
 DB 177 NSLKKHSHQDRVSRSSPLKGTGTPPHEHICLSLPELPGQSPVPTQGRSTSDSPCR 236  
 QY 240 STATQMAPPGPPAAPGGRGSHRDIHQADVRLEATEEIIYLTVPQPPDAEPTSAF 299  
 DB 237 TAATQMAPPGPPAAPGGRGSHRDIHQADVRLEATEEIIYLTVPQPPDAEPTSAF 296  
 QY 300 LPPTESRMSVSSDDPRAAYSTAGRPHPSISBEEGDFCLSSPERAEPGGGWRGSLGP 359  
 DB 297 LPPTESRMSVSSDDPRAAYSVTAGRPHPSISEEDGDFCLSSPERAEPGGGWRGSLGP 356  
 QY 360 PPPRASLSSTALSVDYKTYLVDEHAQLVLSLPCFGYSDSDSATVYDNCASV 419  
 DB 357 PPPRASLSSTALSVDYKTYLVDEHAQLVLSLPCFGYSDSDSATVYDNCASA 416  
 QY 420 SSPESATGEYEAPRPPACLSDESTDEPDVHFKFLNFMVSGRSRSSAESFGL 479  
 DB 417 SSPESALGEYEAPRPPACLSDESTDEPDVHFKFLNFMVSGRSRSSAESFGL 476  
 QY 480 FSCINGEQQTHRAIFRVPVPRHEDELEVDPLLVLEQAEDYIYAYNMTGARGVF 539  
 DB 477 FSCVINGEQQTHRAIFRVPVPRHEDELEVDPLLVLEQAEDYIYAYNMTGARGVF 536  
 QY 540 PAYIAEVTKEPEHMAALAKNSDWDQFRVFKLGSVQVYHKGNVLCRAMQKIATTRL 599  
 DB 537 PAYIAEVTKEPEHMAALAKNSDWDQFRVFKLGSVQVYHKGNVLCRAMQKIATTRL 596  
 QY 600 TVHFNPPSSCVLEISRGVKGIVKADDSQBAKNGKCSHFQKLNISFCGYHPKNNKYFG 659  
 DB 597 TVHFNPPSSCVLEISRGVKGIVKADDSQBAKNGKCSHFQKLNISFCGYHPKNNKYFG 656  
 QY 660 ITKHPADHRFACHFVSEDSTKALAESVGRFAFOQFYKQFVEYTCPTEDIYLE 711  
 DB 657 ITKHPADHRFACHFVSEDSTKALAESVGRFAFOQFYKQFVEYTCPTEDIYLE 708

## RESULT 4

JIP2\_MOUSE  
 ID JIP2\_MOUSE STANDARD; PRT; 830 AA.  
 AC Q9RE9; Q9CX14;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE C-Jun-amino-terminal kinase interacting protein 2 (JNK-interacting  
 DE protein 2) (JIP-2) (JNK MAP kinase scaffold protein 2) (Islet-brain-2)  
 DE (IB-2) (Mitogen-activated protein kinase 8-interacting protein 2).  
 GN MAPK8IP2 OR JIP2 OR IB2.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RC SEQUENCE FROM N.A., AND INTERACTION WITH APOER2.  
 RP STRAIN=BALB/c; TISSUE=Brain;  
 RX MEDLINE=20400499; PubMed=10827199;  
 RA Stockinger W., Brandes C., Fasching D., Hermann M., Gotthardt M.,  
 Herz J., Schneider W.J., Nimpf J.;  
 FT "The reelin receptor ApoER2 recruits JNK-interacting proteins-1 and  
 FT -2";  
 RL J. Biol. Chem. 275:25625-25632(2000).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Embryonic head;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.F.,  
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 Hayashizaki Y.;  
 RL "Functional annotation of a full-length mouse cDNA collection";  
 RT Nature 409:685-690(2001).  
 CC -!- FUNCTION: The JNK-interacting protein (JIP) group of scaffold  
 CC proteins selectively mediates JNK signaling by aggregating  
 CC specific components of the MAPK cascade to form a functional JNK  
 CC signaling module. JIP2 inhibits IIL beta-induced apoptosis in  
 CC insulin-secreting cells (By similarity).  
 CC -!- SUBUNIT: Forms homo- or heterooligomeric complexes. Binds specific  
 CC components of the JNK signaling pathway namely JNK, MAPK7 and  
 CC MLK2, MLK3 and DLK (By similarity). Also binds the proline-rich  
 CC domain-containing splice variant of apolipoprotein E receptor 2  
 CC (ApoER2). Binds the TPR motif-containing C-terminal of kinesin  
 CC light chain. Binds the cytoplasmic tails of LRP1 and LRP2  
 CC (Megalin).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Accumulates in cell surface  
 CC projections (By similarity).  
 CC -!- TISSUE SPECIFICITY: Highly expressed in brain. Expressed in all  
 CC neurons. Also expressed in testis, primarily in the epididymal  
 CC epididymis.  
 CC -!- INDUCTION: Upon neuron differentiation.  
 CC -!- SIMILARITY: BELONGS TO THE JIP SCAFFOLD PROTEINS FAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 PID DOMAIN.  
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to a stop  
 CC codon in position 701.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF310135; AAC31800.1;  
 CC EMBL: AK014339; BAB29282.1; ALT\_SEQ.  
 CC MGD: MGI:1926555; Mapk8ip2.  
 CC InterPro: IPR000050; PID\_domain.  
 DR DR  
 DR DR





Db 598 CLVNGEEREQTHRAVFRFIPRHHPPDELELDVDPVLEAEEDDFWFRGNMRTGERGVPPA 657  
 QY 542 YVALEVTKEPEHMAALAKNSDWQDQVRVFLGSVQVPHYKGNVLCVCAAMOKIATRRLLTV 601  
 Db 658 FYAHAVPGPKDLGSKRSPCWVERFDVFLGSVEVPHCGNGILCAAMOKIATARKLTV 717  
 QY 602 HFNPPSCVLEISVRGVKIGVAKKADDSQEAQKNGKSHFFQKLNISFCGYPHKNKYFGFIT 661  
 Db 718 HLRPPASCDLEISLRGVKLSLGGGP-----FQCSHFFQKLNISFCGYPHKNKYFGFIT 774  
 QY 662 KHPADHFAFCHVFSBDSKALAESVGRFAQFQYKQFVEYTCPTEDIYLE 711  
 Db 775 KHPLLSFACHVFSQSMRPAQSVGRFALEYQYHIACTEDILE 824  
 RESULT 6  
 SP51\_DROME STANDARD; PRT; 490 AA.  
 AC Q9W0K0; Q9NH69;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Eye developmental protein SP512.  
 GN SP512 OR CG1200.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Serano T.L., Pendleton J.D., Rubin G.M.;  
 RT "A reverse genetic screen for genes involved in Drosophila eye  
 RL development";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananietis P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Fandell M.D., Zhang Q., Chen L.X.,  
 RA Stanton R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieux E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Folsler K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegian C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Poirard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen K.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -!- TISSUE SPECIFICITY: Expressed in the embryonic CNS and PNS and  
 CC posterior to the morphogenetic furrow in the eye imaginal disk.  
 CC -!- SIMILARITY: BELONGS TO THE JIP SCAFFOLD PROTEINS FAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 PID DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF231037; AAF34806.1; -;  
 DR EMBL: AE003470; AAF47446.1; -;  
 DR FlyBase: FBgn0040281; SP512.  
 DR InterPro: IPR000050; PID\_domain.  
 DR InterPro: IPR001452; SH3.  
 DR Pfam: PF00018; SH3; 2.  
 DR Pfam: PF00640; PID; 2.  
 DR ProDom: PD000066; SH3; 1.  
 DR SMART: SM00462; PTB; 1.  
 DR SMART: SM00326; SH3; 1.  
 DR SMART: PS01179; PID; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 KW SH3 domain; Developmental protein.  
 FT DOMAIN 150 153 POLY-ARG.  
 FT DOMAIN 185 189 POLY-GLY.  
 FT DOMAIN 251 260 PRO-RICH.  
 FT DOMAIN 271 333 SH3.  
 FT DOMAIN 344 479 PID.  
 FT CONFLICT 164 171 SSILHLG -> C (IN REF. 2).  
 SQ SEQUENCE 490 AA; 53815 MW; 2744EB5790EE043F CRC64;  
 Query Match 14.2%; Score 537.5; DB 1; Length 490;  
 Best Local Similarity 39.4%; Pred. No. 1.3e-20;  
 Matches 114; Conservative 60; Mismatches 86; Indels 29; Gaps 8;  
 QY 443 LSEDSITPDPF-----DV---HFSKKFLNFMNMGSRSSSAESFGLFSCIING-----E 487  
 Db 211 MBDSDSPDSERMQSLGDVDSGHSTAHSPNDFKMSPTQTSVPVSQSPRPFGVPGQLE 270  
 QY 488 EQBQTHRAIFRVPVPRHEDELEVDLPVLVEQADYWEAYNMTGARGVFPAYAYEV 547  
 Db 271 MDEATHRGHLHKFVPRHHDIELEIGDAIYVOKEADLMCEGVNLTGRQGIFFSAYAVDL 330  
 QY 548 T-KEPEHMAALAKNSDWQDQVRVFLGSVQVPHYKGNVLCVCAAMOKIATRRLLTVHFNPP 606  
 Db 331 DYNEFPTVOLVKK-----ERYLLGLVGSVETLAHKGTVGVCQAVRKI-----VGEYGNSP 381  
 QY 607 S--SCVLEISVRGVKIGVKYKADDSQEAQKNGK--CSHFFQKLNISFCGYPHKNKYFGFIT 662  
 Db 382 TGTGTCLEVSDQGLRVDNRSGPNQNKDKKPCIDIFYSLKNVSCFAPHRDHFIFGITK 441  
 QY 663 HPADHFAFCHVFSBDSKALAESVGRFAQFQYKQFVEYTCPTEDIYLE 711  
 Db 442 HPTVQRFACHVFKGSESTRPVAEAVGRAFORFYQKFIETAYPIEDIYIE 490  
 RESULT 7  
 SYNL\_HUMAN STANDARD; PRT; 705 AA.  
 ID SYNL\_HUMAN  
 AC P17600; O75825;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Synapsin I (Brain protein 4.1).



RP SEQUENCE OF 1-1860 FROM N.A.  
RX MEDLINE=93272029; PubMed=8499947;  
RA Iris F.J.M., Bougueret L., Prieur S., Caterina D., Primas G.,  
RA Perrot V., Jurka J., Rodriguez-Tome P., Claverie J.-M., Dausset J.,  
RA Cohen D.;  
RT "Dense Ali clustering and a potential new member of the NF kappa B  
family within a 90 kilobase HLA class III segment.";  
RL Nat. Genet. 3:137-145(1993).  
CC -!- FUNCTION: UNKNOWN.  
CC -!- TISSUE SPECIFICITY: LIMITED TO CELL-LINES OF LEUKEMIC ORIGIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; M33509; AAA35585.1; -;  
DR EMBL; M33518; AAA35586.1; -;  
DR EMBL; M33512; AAA35586.1; JOINED.  
DR EMBL; M15025; AAA78744.1; -;  
DR PIR; B35098; B35098.  
DR PIR; S36152; S36152.  
DR Genew; HGNC:13918; BAT2.  
DR MIM; 142580; -;  
KW Repeat.  
FT DOMAIN 519 524 POLY-RICH.  
FT GLN-RICH.  
FT 636 657  
FT POLY-PRO.  
FT 684 688  
FT POLY-PRO.  
FT 699 704  
FT POLY-PRO.  
FT 814 821  
FT POLY-PRO.  
FT 1340 1345  
FT POLY-GLY.  
FT 1398 1403  
FT POLY-GLY.  
FT 1436 1442  
FT POLY-PRO.  
FT 1982 1991  
FT POLY-PRO.  
FT 41 1795  
FT 41 95  
FT REPEAT 41 95  
FT 98 154  
FT REPEAT 98 154  
FT 281 337  
FT REPEAT 281 337  
FT 1740 1795  
FT REPEAT 1740 1795  
FT 337 549  
FT REPEAT 337 549  
FT 476 549  
FT REPEAT 476 549  
FT 1899 2089  
FT REPEAT 1899 2089  
FT 1899 1948  
FT REPEAT 1899 1948  
FT 1965 2014  
FT REPEAT 1965 2014  
FT 2040 2089  
FT REPEAT 2040 2089  
FT 57 57  
FT CONFLICT 57 57  
FT 109 109  
FT CONFLICT 109 109  
FT 414 414  
FT CONFLICT 414 414  
FT 532 532  
FT CONFLICT 532 532  
FT 682 682  
FT CONFLICT 682 682  
FT 730 730  
FT CONFLICT 730 730  
FT 750 750  
FT CONFLICT 750 750  
FT 834 834  
FT CONFLICT 834 834  
FT 1035 1035  
FT CONFLICT 1035 1035  
FT 1068 1068  
FT CONFLICT 1068 1068  
FT 1285 1285  
FT CONFLICT 1285 1285  
FT 1400 1400  
FT CONFLICT 1400 1400  
FT 1611 1611  
FT CONFLICT 1611 1611  
FT 1729 1729  
FT CONFLICT 1729 1729  
FT 2142 AA; 227840 MW; 32DF16B9B52420A CRC64;  
SQ SEQUENCE 2142 AA; 227840 MW; 32DF16B9B52420A CRC64;  
Query Match 5.1%; Score 194.5; DB 1; Length 2142;  
Best Local Similarity 23.2%; Pred. No. 0.012;  
Matches 118; Conservative 33; Mismatches 170; Indels 187; Gaps 22;  
QY 13 AASPRAAPFLGHIAFPNFRTHDLSLEFPD-EDLSEITDECGISLQCKDXTLSLRPP 71  
DB 1281 APAPRAAA-----AKSPDLNSNSDOANEWEETASESSDFTSERGDKAEPPVLLTPK 1334

QY 72 RAGLLSAGGGAGSRLQAEMLQMDLIDATGTPGAEDDEDDDEARRRCPGAPPKAES 131  
DB 1335 AVTPOCGGGGAVPGISA-MSRGDLISQRAKDLISKRSFSQSPGMRNRRPCGKAGSS 1393  
QY 132 GORPASRGQGGQSGQSGQSGDTPYRKPTTLNLFQVPRSQDTLNNNSLGKHKHSWDR 191  
DB 1394 GS--SSGGGGGGGGTGPGRGD-----KRSWSP 1421  
QY 192 VSRSSSPLTKGTQTPPHEHICLSEELPPSGPAPTDRCTSTDSFCRSTA----- 242  
DB 1422 KNSRPP-----EERPP-----GLPLPP-----PPSSAVFRLDQVHSPAGIQALQAL 1468  
QY 243 -----TOMAPPG-----GPPAAPGCGRSHRDRTHYQADVRLTEATEEIIYLPVQRPDP 291  
DB 1469 SSRQGSVTAPGGHPRHKPGPPQAPG-----PSRPPET 1501  
QY 292 AAPTSAFLPPTESRMSVSSDP-----DPAAY-----PSTAGRP--HPS 328  
DB 1502 RYEPQRV-----NSGLSSDPHFEEPPGMVRGVGTGPRDSAGVSPFPKRRERPPKPE 1554  
QY 329 ISEEEGFDCLSSPERAEPGGGWRGSLGEPPPPRLASLSSDTSALS-----YDSVKYTLVV 385  
DB 1555 LLOEE-----SLP-----PHSSGFLGSKGEGPQQAESRDTGTTEALTPHWNRL----- 1599  
QY 386 DEHAQLELVSLRPFCDGSDSATVDNCASVSSPSYSAIEEYEEAPRPOPPACLSE 445  
DB 1600 --HTATSRKSYRP-----TSMBPWMEPLSPFFEDVAGTMSQS----- 1634  
QY 446 DSTPDEPDVHFHFKFLNLFVMSGRSRSS 473  
DB 1635 DSGVD-----LSGDSQVSS 1648  
RESULT 9  
WAIP\_HUMAN STANDARD; PRT; 503 AA.  
ID AC 043516; Q9UNP1; Q15220;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Wiskott-Aldrich syndrome protein interacting protein (WASP interacting  
DE protein) (PRPL-2 protein).  
GN WASPIP OR WIP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98070810; PubMed=9405671;  
RA Ramesh N., Anton I.M., Hartwig J.H., Geha R.S.;  
RT "WIP, a protein associated with Wiskott-Aldrich syndrome protein,  
RT induces actin polymerization and redistribution in lymphoid cells.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:14671-14676(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Tonsil;  
RA Kreideweiss S., Delany-Heiken P., Nordheim A., Ruhlmann A.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 192-503 FROM N.A., AND VARIANT ALA-495.  
RX MEDLINE=99218549; PubMed=10202051;  
RA Stewart D.M., Tian L., Nelson D.L.;  
RT "Mutations that cause the Wiskott-Aldrich syndrome impair the  
RT interaction of Wiskott-Aldrich syndrome protein (WASP) with WASP  
RT interacting protein.";  
RL J. Immunol. 162:5019-5024(1999).  
CC -!- FUNCTION: MAY HAVE DIRECT ACTIVITY ON THE ACTIN CYTOSKELETON.  
CC INDUCES ACTIN POLYMERIZATION AND REDISTRIBUTION.  
CC -!- SUBUNIT: BINDS TO WASP, PROFILIN AND ACTIN.  
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PERIPHERAL BLOOD  
CC MONONUCLEAR CELLS. SPLEEN, PLACENTA, SMALL INTESTIN, COLON,  
CC THYMUS. LOWER EXPRESSION IN OVARY, HEART, BRAIN, LUNG, LIVER, CC







```
SQ SEQUENCE 1744 AA: 187214 MW: 5C3C8B6211935524 CRC64;
Query Match 4.7%; Score 179; DB 1; Length 1744;
Best Local Similarity 19.8%; Pred. No. 0.058;
Matches 177; Conservative 312; Mismatches 312; Indels 306; Gaps 39;
QY 13 AASPPAASFFGLHIASPN-----FLRTHDISLEFEDEDLSEITDEGICISLQCKDITL 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 928 ASYSPAGSQ--QLLVSSPPSTAPAQSQLPHK-GLESYED-----LSRSGEPL 973
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 67 SLRPPRAGLLSAGGGAGSRLOAEMLOMLDITADGTFGAEDDEDDEREAARRP----- 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 974 NLE-----GLAVRHVAGVQSRKSP--EESTVPARRRTPSDSHYEKSSPEPGSPRSPTVLS 1027
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 123 -----GAGPPK-----AESGQEPASRGQOGQOGQSGDGT-----YR 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1028 PEVYSTIAANPGGRKEPHLSHYEAEMEASPSLSLTSGGVRSPPGLAKTPLSALGLK 1087
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 158 PKRPTTLNLF-----QVPRS-----QDTLNN 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1088 PHNPADILLHPVGELEGEAGADSEEPSRYVESVARTATTGRAGNLPAAPVGLVPAARN 1147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 NSLGGKHSWQDVRSSSPLKT-----GEQTPPHEHICLSEELPPQSGP 223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1148 GAFGNSFTVPSPVS-TSPITHSVDCASLRYSPESGSPHGTVPHPA-VAETAYRSPMVQ 1205
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 224 APTDRGTSTDSP-----MSAGERS-----NSLPNATVNGKASPLSGMS 1429
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1206 TPAHSSVQTSPPSFOAGTLGSPYASPDYDGRGGFQDPQARQOPQVSVGVHALPGS 1265
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 ---CRRSTATOMAP-----PGGP-----PAAPGGRGSHRDRIHY 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1266 PRTLHRTVATNTPSPGFRRAANPAVASVPGSPGLGRHTVSPHAPGPSLAR----HQ 1321
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 270 QADYRLATEIYLTVPQPPDAEPTSA--FLPPTESMSV-----SSDDP 315
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1322 MAAPPGSPMVGYSPEERAPTLRSQSSASGYQPPSPFVSPAYPGTSTHSSSPDS 1381
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 316 AAYPSTAGRPHPSISEEEGFCILSSPERAEPPGGGWRGSL-----GPPPPPPRASLS 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1382 AAYRQGSPTOPALPEKRR-----MSAGERS-----NSLPNATVNGKASPLSGMS 1429
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 369 SDTSALSYDSVKYTLVDEHAQLVSLRPFCDYSDSDSATVYDNCASVSSP-----YE 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1430 SPSSG-SAVAFSHTL-----PDFSKFSPDISPETRANVFKVQDTSKYWK 1474
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 425 SAICEEVEEAPRPPACLSQSTDEPDVHFSSKFLNVFMSGRSSSSAESGLFSCII 484
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1475 PDISREQAIA-----LLKDRPGAFIIRDSHSFRGAYGLAMKVASPPP-----TVMQ 1521
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 485 NGEQEQTHRAIRFV-----PRHEDELELEVDDPLLVQLAEQDYWEAYNMRTGARGVFP 540
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1522 QNKKGDITNELVRHFLTETSPRGV-KLKGCPNEPFGCLLSALYQHSIMPLALPCKLVIP 1580
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 541 AYYAIEVTKPEHMAALAKNS--DWDDQ----FRVKFLGVSQVYHKGNDVLCVAAQMKI-- 593
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1581 DRDPMEKED-----AASSTNSATDLKQGAACNVLFINSVEMESLGTGPOAISKAVATILV 1636
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 594 --ATTRLRLTVHFNPPSCVLEISVRGVKIGVKADDQSEAKGKCSHFQ----LKNISFC 647
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1637 ADPTPTATIVHF-----KVAQGITL-----TDNQKULFRFRHPLNVTVTFC 1678
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 648 GYHPKNNKY-----FGFIT-KHPADHRFACHVFVSDSTKALAESV 687
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1679 DLDPOERKWKTKDGSFPAKLFGEVARKQGSTTDNVCHLFAELDPDQPAAIIV 1730
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 12
IRSI_HUMAN
ID IRSI_HUMAN STANDARD; PRT; 1242 AA.
AC P35568;
DC 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
Insulin receptor substrate-1 (IRS-1).
IRSI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=93292738; PubMed=8513971;
RA Araki E., Sun X.J., Haag B.L. III, Chuang L.M., Zhang Y.,
RA Yang-Feng T.L., White M.F., Kahn C.R.;
RT "Human skeletal muscle insulin receptor substrate-1. Characterization
of the cDNA, gene, and chromosomal localization.";
RL Diabetes 42:1041-1054(1993).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=92181456; PubMed=13111924;
RA Nishiyama M., Wands J.R.;
RT "Cloning and increased expression of an insulin receptor substrate-1-
like gene in human hepatocellular carcinoma.";
RL Biochem. Biophys. Res. Commun. 183:280-285(1992).
RN [3]
VARIANTS PRO-512 AND ARG-971
RX MEDLINE=93390176; PubMed=8104271;
RA Almind K., Bjoergbaek C., Vestergaard H., Hansen T., Echwald S.,
RA Pedersen O.;
RT "Aminoacid polymorphisms of insulin receptor substrate-1 in
non-insulin-dependent diabetes mellitus.";
RL Lancet 342:828-832(1993).
RN [4]
VARIANT NIDDM GLY-723 DEL.
RX MEDLINE=96303710; PubMed=8723689;
RA Esposito D.L., Mammarella S., Ranieri A., della Loggia F., Capani F.,
RA Consoli A., Mariani-Costantini R., Caramia F.G., Cama A., Battista P.;
RT "Deletion of Gly723 in the insulin receptor substrate-1 of a patient
with noninsulin-dependent diabetes mellitus.";
RL Hum. Mutat. 7:364-366(1996).
RN [5]
VARIANTS NIDMM TYR-1043 AND TYR-1095.
RA Mammarella S., Creati B., Esposito D.L., Arcuri P., della Loggia F.,
RA Capani F., Mariani-Costantini R., Caramia F.G., Battista P., Cama A.;
RT "Novel allele of the insulin receptor substrate-1 bearing two
non-conservative amino acid substitutions in a patient with
noninsulin-dependent diabetes mellitus.";
RL Hum. Mutat. 11:411-411(1998).
RN [6]
STRUCTURE BY NMR OF 157-267.
RX MEDLINE=96185451; PubMed=8599766;
RA Zhou M.-M., Huang B., Olejniczak E.T., Meadows R.P., Shuker S.B.,
RA Miyazaki M., Trueb T., Shoelson S.E., Fesik S.W.;
RT "Structural basis for IL-4 receptor phosphopeptide recognition by the
IRS-1 PTB domain.";
RL Nat. Struct. Biol. 3:388-393(1996).
CC !- FUNCTION: MAY MEDIATE THE CONTROL OF VARIOUS CELLULAR PROCESSES
BY INSULIN. WHEN PHOSPHORYLATED BY THE INSULIN RECEPTOR BINDS
SPECIFICALLY TO VARIOUS CELLULAR PROTEINS CONTAINING SH2 DOMAINS
SUCH AS PHOSPHATIDYLINOSITOL 3-KINASE P85 SUBUNIT OR GRB-2.
CC !- DISEASE: POLYMORPHISMS IN IRS1 MAY BE INVOLVED IN THE ETIOLOGY OF
A SUBSET OF LATE-ONSET NON-INSULIN-DEPENDENT DIABETES MELLITUS
(NIDDM).
CC !- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC !- SIMILARITY: CONTAINS 1 PTB DOMAIN.
CC
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
```





Search completed: December 24, 2002, 08:49:40  
Job time : 46 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2002, 08:48:50 ; Search time 101 Seconds  
(without alignments)  
1450.490 Million cell updates/sec

Title: US-09-966-561-2

Perfect score: 3792

Sequence: 1 MAERESGGLGGAASPPAAS.....QQFYKQVEYTCPTEDIVLE 711

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.21.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phage.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3541	93.4	707	11 Q925J8	Q925J8 mus musculus
2	1100	29.0	455	4 Q96G62	Q96G62 homo sapien
3	246.5	6.5	315	5 P91045	P91045 caenorhabdi
4	211	5.6	59	11 Q9CUX3	Q9CUX3 mus musculus
5	210	5.5	647	10 Q8S148	Q8S148 oryza sativ
6	210	5.5	950	11 Q923A9	Q923A9 mus musculus
7	210	5.5	2157	11 Q9Z1R1	Q9Z1R1 mus musculus
8	200.5	5.3	1217	4 Q9ULL5	Q9ULL5 homo sapien
9	190.5	5.0	1735	4 Q9HBL0	Q9HBL0 homo sapien
10	190	5.0	2157	4 Q95875	Q95875 homo sapien
11	190	5.0	2157	4 Q96QC6	Q96QC6 homo sapien
12	184	4.9	1715	6 Q9GLM4	Q9GLM4 bos taurus
13	183.5	4.8	640	4 Q60593	Q60593 homo sapien
14	182	4.8	1634	5 Q9XW25	Q9XW25 caenorhabdi
15	180.5	4.8	1251	6 Q28224	Q28224 cercopithe
16	179.5	4.7	4957	4 Q14687	Q14687 homo sapien

17	179.5	4.7	5262	4	Q14686	Q14686 homo sapien
18	179	4.7	623	5	Q9W4A2	Q9W4A2 drosophila
19	178	4.7	926	5	Q9W3G1	Q9W3G1 drosophila
20	177.5	4.7	1240	13	P79773	P79773 gallus gall
21	177	4.7	1240	12	Q9DMH8	Q9DMH8 rat cytoleg
22	176.5	4.7	917	4	Q9PIY5	Q9PIY5 homo sapien
23	173.5	4.6	645	4	Q96EX0	Q96EX0 homo sapien
24	173.5	4.6	1110	5	Q09493	Q09493 caenorhabdi
25	173	4.6	304	4	Q9UBP9	Q9UBP9 homo sapien
26	172	4.5	857	3	Q9P7E8	Q9P7E8 schizosacch
27	171.5	4.5	619	5	Q960W5	Q960W5 drosophila
28	171.5	4.5	1077	5	Q9VDK1	Q9VDK1 drosophila
29	171.5	4.5	1279	5	Q97005	Q97005 leishmania
30	171.5	4.5	1307	10	Q9LVN1	Q9LVN1 arabidopsis
31	171	4.5	302	11	Q9CRV4	Q9CRV4 mus musculus
32	170.5	4.5	440	4	Q8TF74	Q8TF74 homo sapien
33	170.5	4.5	440	4	Q8TE44	Q8TE44 homo sapien
34	170.5	4.5	487	11	Q8VDA4	Q8VDA4 rattus norv
35	170.5	4.5	1056	4	Q9DWM7	Q9DWM7 rat cytoleg
36	170.5	4.5	1182	4	Q95621	Q95621 homo sapien
37	170.5	4.5	1182	4	Q99495	Q99495 homo sapien
38	170.5	4.5	1190	4	Q99621	Q99621 homo sapien
39	170.5	4.5	1201	11	Q8VHK1	Q8VHK1 mus musculus
40	170.5	4.5	2158	11	Q9WU13	Q9WU13 rattus norv
41	170	4.5	309	11	Q9CYD2	Q9CYD2 mus musculus
42	169.5	4.5	1044	4	Q9P206	Q9P206 homo sapien
43	169.5	4.5	2112	5	Q9VEL9	Q9VEL9 drosophila
44	169.5	4.5	3105	12	Q8UZ19	Q8UZ19 cercopithic
45	169	4.5	1386	4	Q9C0A3	Q9C0A3 homo sapien

## ALIGNMENTS

RESULT 1

Q925J8  
ID Q925J8 PRELIMINARY; PRT; 707 AA.  
AC Q925J8;  
DT 01-DEC-2001 (TREMREL. 19, Created)  
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)  
DE 01-JUN-2002 (TREMREL. 21, Last annotation update)  
GN MAPK8IP OR PRK8IP.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IUS;  
RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,  
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;  
RT "High-Throughput Sequence Identification of Gene Coding Variants  
Within Alcohol-Related QTLs";  
RL Submitted (DGC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF332076; AA056104.1; -  
DR MGD; MGI:1309464; Mapk8ip.  
DR InterPro; IPR000050; PID\_domain.  
DR InterPro; IPR001452; SH3.  
DR Pfam; PF00640; PID; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PROSITE; PS01179; PID; 1.  
DR PROSITE; PS00002; SH3; 1.  
KW Kinase.  
SQ SEQUENCE 707 AA; 77381 MW; FDA53D891B6F5AA6 CRC64;

Query Match 93.4%; Score 3541; DB 11; Length 707;

Best Local Similarity 93.5%; Pred. No. 1.3e-250;

Matches 666; Conservative 13; Mismatches 27; Indels 6; Gaps 3;

QY 1 MAERESGGLGGAASPPAASPPFLGLHTASPPNFRLTHTDLSLEEFEDDLSEITDECGISL 60

DB 1 MAERES-GLGRGAASPPAASPPFLGLHTASPPNFRLTHTDLSLEEFEDDLSEITDECGISL 59

```

QY 61 QCKDTLSLRPPRAGLLSAG-GGGAGSRLQAEMLQMDLIDATGTPGAEDDEEDDEERAA 119
Db 60 QCKDTLSLRPPRAGLLSAGSGSAGSLQAEMLQMDLIDAGDTPGAEDDEEDDEERAA 119
QY 120 RPPGAGPPKAESGQEPASRGQSGQSGQSGQSGDYRPRPTTLNLFQVPRSQDTLNN 179
Db 120 QRPVGGPPKAESGQNDPAPR-----SQGGGTGSGDYRPRPTTLNLFQVPRSQDTLNN 175
QY 180 NSLAKKHSWQDRVSRSSPLKGTQTPPHEHICLSELPQSGPAPTDRGTSTDSPCR 239
Db 176 NSLAKKHSWQDRVSRSSPLKGTQTPPHEHICLSELPQSGPAPTDRGTSTDSPCR 235
QY 240 STATOMAPPGPPAPPGCGRSHRDIHYQADVRLATEEIIYLPVQRPDAAEPTSAF 299
Db 236 SAATOMAPPGPPAPPGCGRSHRDIHYQADVRLATEEIIYLPVQRPDAAEPTSTF 295
QY 300 LPPTESRMSVSDPDPAAPSTAGRPHPSISEEEGFCCLSSPERAEPGGGWRGSLGP 359
Db 296 MPPTESRMSVSDPDPAAPSTAGRPHPSISEEEGFCCLSSPERAEPGGGWRGSLGP 355
QY 360 PPPRASLSSTALSVDKYLTVDEHAQLELVSLRCPGFGYSDSESATVYDNCASV 419
Db 356 PPPRASLSSTALSVDKYLTVDEHAQLELVSLRCPGFGYSDSESATVYDNCASA 415
QY 420 SSPYESAIGEYEEAPRPPACLSDESTDPDVFHFSKFLNFMVMSGRSSSAESFGL 479
Db 416 SSPYESAIGEYEEAPRPPACLSDESTDPDVFHFSKFLNFMVMSGRSSSAESFGL 475
QY 480 FSCIINGEQQTURATFRVPRHEDELEVDLPLVLAQADYWEAYNMRTGARGVF 539
Db 476 FSCVINGEQQTURATFRVPRHEDELEVDLPLVLAQADYWEAYNMRTGARGVF 535
QY 540 PAYIAEYVTEKEPHEMAALAKNSDWQDFRVKFLGSGVQVYKHGNDVLCAMOKIATTRL 599
Db 536 PAYIAEYVTEKEPHEMAALAKNSDWQDFRVKFLGSGVQVYKHGNDVLCAMOKIATTRL 595
QY 600 TVHNPPSSCVLEISVRGKIVGKADDSQBAKNGKSHFFQKNIISFCGYHPKNNKYFGF 659
Db 596 TVHNPPSSCVLEISVRGKIVGKADDSQBAKNGKSHFFQKNIISFCGYHPKNNKYFGF 655
QY 660 ITKHPADHRFACHVFVSEDSTKALAESVGRFAQFQKQFVEYTCPTEDIYLE 711
Db 656 ITKHPADHRFACHVFVSEDSTKALAESVGRFAQFQKQFVEYTCPTEDIYLE 707

RESULT 2
Q96G62
ID Q96G62 PRELIMINARY; PRT; 455 AA.
AC Q96G62;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to mitogen-activated protein kinase 8 interacting protein 2 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009940; AA009940.1; -.
DR InterPro; IPR000050; PID_domain.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00640; PID; 1.
DR Pfam; PF00018; SH3; 1.
DR PROSITE; PS01179; PID; 1.
DR PROSITE; PS00002; SH3; 1.
KW Kinase.
FT NON_TER 1 1

```

```

SQ SEQUENCE 455 AA; 50405 MW; F10CD3C4C6A8E522 CRC64;
Query Match 29.0%; Score 1100; DB 4; Length 455;
Best Local Similarity 35.7%; Pred. No. 2.4e-72;
Matches 252; Conservative 64; Mismatches 95; Indels 294; Gaps 14;

QY 26 HIASPPNFRUTHISDEFEDSEITDECGISL-----QC-KDTLSLRPPRAGLLSAG 79
Db 26 HSLSPGCRPPQDLSIEEFDDELSETDDGLGLSDHCKDLSL----- 74
QY 80 GGGAGS-----LQAEMLQMDLIDATGTPGAEDDEEDDEERARRPGAGPPKAESQ 133
Db 75 --GRSEDPHICFQDFQFEMID-----DNEEEDEDEEEDEAE----- 114
QY 134 EPASRGQSGQSGQSGQSGDYRPRKPTTLNLFQVPRSQDTLNNNSLGKHKHSDRVS 193
Db 115 ----- 114
QY 194 RSSPLKGTQTPPHEHICLSELPQSGPAPTDRGTSTDSPCRSTATQMAPGGPPA 253
Db 115 -----DS-----A 117
QY 254 APGGRGHRSHRDIHYQADVRLATEEIIYLPVQRPDAAEPTSAFLPPTESRMSVSDP 313
Db 118 GSPGGRGTG----- 126
QY 314 DPAAYPSTAGRPHPSISEEEGFCCLSSPERAEPGGGWRGSLGEBPPPPRASLSDDTSA 373
Db 127 -----PSAPR-----DAS 134
QY 374 LSVDKYLTVDEHAQLELVSLRCPGFG-DYSDSESATVYDNCASVSPYESAIGEYEE 432
Db 135 LVYDAVKYLTVDEHTQLELVSLRRCAGLGHDSG-----GEASE 177
QY 433 EAPRQPPACL-----SESTDEPDVFHFSKFLNFMVMSGRSSSAESFGLFSCIING 486
Db 178 E-----EAGALLGGQVSGDTPSDPLTFKFLNFMVNSRSTSTESFGLFSCIING 233
QY 487 EEOQTHRAIFRVPRHEDELEVDLPLVLAQADYWEAYNMRTGARGVFPAYIAE 546
Db 234 EEREQTHRAVFEIPIRHPDELDVDPVLEAEEDDFWFRGNMRTGARGVFPAYIAH 293
QY 547 VTKEPEHMAALAKNSDWQDFRVKFLGSGVQVYKHGNDVLCAMOKIATTRLTVHFNPP 606
Db 294 VPGPAKDLGSKRSPCWVERFDVQFLGSEVPCHQNGILCAAMOKIATARKLTVHLRPP 353
QY 607 SSCVLEISVRGKIVGKADDSQBAKNGKSHFFQKNIISFCGYHPKNNKYFGFITKHPAD 666
Db 354 ASCDLEISLRGVKLSLGGGPE--FQRCSHFFQMKNISFCGCHPRNSCYFGFITKHPLL 410
QY 667 HRFACHVFVSEDSTKALAESVGRFAQFQKQFVEYTCPTEDIYLE 711
Db 411 SRFACHVFVSESMPRVAQSVGRFALEYQEHLAYACPTEDIYLE 455

RESULT 3
P91045
ID P91045 PRELIMINARY; PRT; 315 AA.
AC P91045;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CL3A10.3 protein.
GN CL3A10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;

```



```
Db 293 GDVPPSSSDTSPTTPGGGSPS-----TPCDTPPSPSGTSPTT 331
QY 298 ---AFLPPTESRMVSSDPPDPAAYPSTAGRPHPSISEEBEGFDCLSPERAEPGGGWRG 354
Db 332 PGGGYPTPTSVGDVPPASGTSPTTPGGGGYSPSTPCSAAPPSPSGTSTPTTPGGGYSP 391
QY 355 S--LGEPPPPRPSLSDTSALSVDYSKYTLVVDEHAQLVSLRPGFCGSDSDSATV 412
Db 392 STPCNAPPSP-----SSDTSPTT-----PGGGNYPPAPTIGNV 424
QY 413 YNCASVSPYESAIGEYEEAERP-OPPACLSFSDSTPDEP 452
Db 425 PPSFSGTSP--STPGGCGSSSTPCDAPPSPSSDTSPTTP 463

RESULT 6
Q923A9
ID Q923A9 PRELIMINARY; PRT; 950 AA.
AC Q923A9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN cDNA 3110039B05 gene (Fragment).
GN BAT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006664; AAH06664.1; -.
DR MGD; MGI:1915467; Bat2.
FT NON_TER 1
SQ SEQUENCE 950 AA; 99726 MW; E41DE9AB4A6D976 CRC64;

Query Match 5.5%; Score 210; DB 11; Length 950;
Best Local Similarity 22.9%; Pred. No. 7.7e-07;
Matches 123; Conservative 33; Mismatches 200; Indels 182; Gaps 21;

QY 3 ERESGLGGG-----AASPPAASPFLGLHIASPPNRL---THDISLEEFED---- 46
Db 50 ERENAARGADGKPPSLTLAASFGPEETLTAATVPPPPRTAAKSPDLNSQNSDQANEW 109
QY 47 EDLSEITDECGISLQCKDTLSLRPPRA-----GLLSAGGGGAGSRLOAEMQLMDLIDA 99
Db 110 ETASESSDFASERRGDKET-----PPAALMTSKAVGTPGANAGAGPGISA-MSRGLDSOR 164
QY 100 TGDTPCAEDDEEDDEERAARRPGAGPPKAESGOEPASRGOGSOGSOGSGDITYRKP 159
Db 165 AKDLSKRFSSQRPQMDRQNRRTGCKTGSGG--GSSGGGAGGCGRTGPGRGD----- 217
QY 160 RPTTLMLFQVPRSQDTLNNLSLGGKHSQDRVSRSSSLPKTGEQTPPHEHICLSBELPP 219
Db 218 -----PPQAPQSGSSPRPTRYDPPRAS--SAISSDPHFEEFGPMVRGV 242
QY 220 QSGPAPTTDRGTSTDSPCRSTA-----TQMAPGCGPPAAPPGGRGHSRDR 266
Db 243 ---PPSSSAVFLDQVHNSNFIAGIQQAQLSSRQGNVTAPGGHPRPKPG----- 290
QY 267 IHQADVRLATEEIIYLTVPQRPDAA-EPTSAFLPPTESRMVSSSDP---DPA-AYPST 321
Db 291 -----PPQAPQSGSSPRPTRYDPPRAS--SAISSDPHFEEFGPMVRGV 331
QY 322 AGRPHPSISEEBEGFDCLSPERAERP-----GGGWRGSLGEPGPPPPPRAS 366
Db 332 GGTPRDS-----AGVNPFPKRRERPPKRPPELLQETVPASHSSGFLGSKPEVPGPQES 386
QY 367 LSSDTSALSVDYSKYTLVVDEHAQLVSLRPGF-----GDYSDSDSATVYDNC 416
Db 387 RDSGTEALPTHWNLRLHTATSRKSYQPSIEPMEPLSPFEDVAGTEMQSQSDGVDLSGD 446
```

```
QY 417 ASVSS-----PVESAIGEYEEAERPQPPA 441
Db 447 SQVSSGFCSSORSRSPDGLKSAEGPPKRPQGPPLNAVPGESASGSEPPRRRPPA 504

RESULT 7
Q9Z1R1
ID Q9Z1R1 PRELIMINARY; PRT; 2157 AA.
AC Q9Z1R1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BAT2.
GN NFKB1L1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Rowen L., Qin S., Madan A., Abbasi N., James R., Dickhoff R.,
RA Shaffer T., Ratcliffe A., Loretz C., Lasky S., Hood L.;
RT "Sequence of the mouse major histocompatibility class III region.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF109719; AAC82480.1; -.
DR MGD; MGI:1340031; Nfkb1l1.
SQ SEQUENCE 2157 AA; 229072 MW; 35B21F6B7C971F40 CRC64;

Query Match 5.5%; Score 210; DB 11; Length 2157;
Best Local Similarity 22.9%; Pred. No. 2.3e-06;
Matches 123; Conservative 33; Mismatches 200; Indels 182; Gaps 21;

QY 3 ERESGLGGG-----AASPPAASPFLGLHIASPPNRL---THDISLEEFED---- 46
Db 1258 ERENAARGADGKPPSLTLAASFGPEETLTAATVPPPPRTAAKSPDLNSQNSDQANEW 1317
QY 47 EDLSEITDECGISLQCKDTLSLRPPRA-----GLLSAGGGGAGSRLOAEMQLMDLIDA 99
Db 1318 ETASESSDFASERRGDKET-----PPAALMTSKAVGTPGANAGAGPGISA-MSRGLDSOR 1372
QY 100 TGDTPCAEDDEEDDEERAARRPGAGPPKAESGOEPASRGOGSOGSOGSGDITYRKP 159
Db 1373 AKDLSKRFSSQRPQMDRQNRRTGCKTGSGG--GSSGGGAGGCGRTGPGRGD----- 1425
QY 160 RPTTLMLFQVPRSQDTLNNLSLGGKHSQDRVSRSSSLPKTGEQTPPHEHICLSBELPP 219
Db 1426 -----KRWSPSPKNRSRPP-----EERPP-----GLPLPP 1450
QY 220 QSGPAPTTDRGTSTDSPCRSTA-----TQMAPGCGPPAAPPGGRGHSRDR 266
Db 1451 ---PPSSSAVFLDQVHNSNFIAGIQQAQLSSRQGNVTAPGGHPRPKPG----- 1498
QY 267 IHQADVRLATEEIIYLTVPQRPDAA-EPTSAFLPPTESRMVSSSDP---DPA-AYPST 321
Db 1499 -----PPQAPQSGSSPRPTRYDPPRAS--SAISSDPHFEEFGPMVRGV 1539
QY 322 AGRPHPSISEEBEGFDCLSPERAERP-----GGGWRGSLGEPGPPPPPRAS 366
Db 1540 GGTPRDS-----AGVNPFPKRRERPPKRPPELLQETVPASHSSGFLGSKPEVPGPQES 1594
QY 367 LSSDTSALSVDYSKYTLVVDEHAQLVSLRPGF-----GDYSDSDSATVYDNC 416
Db 1595 RDSGTEALPTHWNLRLHTATSRKSYQPSIEPMEPLSPFEDVAGTEMQSQSDGVDLSGD 1654
QY 417 ASVSS-----PVESAIGEYEEAERPQPPA 441
Db 1655 SQVSSGFCSSORSRSPDGLKSAEGPPKRPQGPPLNAVPGESASGSEPPRRRPPA 1712

RESULT 8
Q9ULL5
ID Q9ULL5 PRELIMINARY; PRT; 1217 AA.
AC Q9ULL5;
```



```
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE KIAA1205 protein (Fragment).
GN KIAA1205.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP Tensin.
RC SEQUENCE FROM N.A.
RC TISSUE=BRNIN;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K., Kikuno R., Hirose M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 6:337-345(1999).
DR EMBL; AB033031; BAA86519.1; -.
DR InterPro; IPR000637; AT_hook.
DR PRINTS; PR002965; P_rich_extensn.
DR PRINTS; PR00929; ATHOOK.
DR PRINTS; PR01217; PRICHEXTENS.
FT NON_TER 1
FT SEQUENCE 1217 AA; 130189 MW; EC83FC3173861B73 CRC64;

Query Match 5.3%; Score 200.5; DB 4; Length 1217;
Best Local Similarity 21.3%; Pred. No. 5.3e-06;
Matches 132; Conservative 57; Mismatches 217; Indels 213; Gaps 26;

QY 11 GGAASP-----PAASFFLGLHIASPNFRLTHDISLEEF-----DEDLSEITDECG 57
DB 238 GGLTSPIFCSTKPKKLLKTSFHLRRRDPF-FQTPKLYAQEYFEADEKADVPAD-- 294
QY 58 ISLCKDTLSLRPPR-AGLLSAGGGAGSLQAEMLQMDLIDATGTPGAEDDEEDDEE 116
DB 295 -----IRLNRLPLDVSS-----CSRPAISPLGIDFCPPNPFGDGRRRGRKP 340
QY 117 RAARRP-----AGPKAESGQEPAS--RGOGOGOGOGOGSGSDTY 156
DB 341 TKAKRDGPPRRGRPRIRPLEVPTTAGPASASTPDCAKKPRGRGRGRKAEAGGTRL 400
QY 157 RPKRPTLNLFPQVPRSDTLNNNSLGKKHMQDRVSRSS-----P- 198
DB 401 EPLKPLKIKL--SVPKAGEGLGTSS-GDAISGTDHNSLDSLTREKIEAKIKEVEEKQPE 457
QY 199 -----LKTGQTPPHEHICLSEEL-PPQSGPAPTTDRGTSTDSPCRSTATOM 245
DB 458 MKSGFMAFSLDFLKGRHHPPLYOAGLTPLSPPKSVPPSVPARGLQPQPPA--TPAVPH 515
QY 246 APPGPP-----PAAPPGGGR-----HSHRDRIHYQADV 273
DB 516 PPSGAFGLGGALEAAESEGILGCPSPCKRLDEELKRNLETLPFSFDEEESVAKNRDL 575
QY 274 RLEATEIYL---TPVORPDAAEPTSAFLPPTESRMVSSDDPAAYPS--TAGRPHPS 328
DB 576 QESISSAISALDPLPAGPKDTSTPDG---PPLAPAAAAPGPPPLGLPLSANSNGTPEPP 632
QY 329 ISSEEEGDFCLSSPERAEPGGGWRGLSGELPPPPPRASLSSDTSALSYSVXYTLVVDH 388
DB 633 LLEEK-----PPPTPPPAPTQPPQPPPPPPPPQPALPSP----- 667
QY 389 AOELVSLRCPFGDYSDSDSATVYDNCASVSSPYESAIGEYEEA-----PRPQPACLS 444
DB 668 ---PLVAPT-----SSPPPPPLPPPPPPPPPPPPPPPPPPPPPPPPPPPP 703
QY 445 EDSTPDEP-----DVHFSKFLNVFMGSRSSSAESFGLFCIIINGEEOE 490
DB 704 LAAPPEEPAAPSPEDPELPDTRPLHLAKQETAACVGETDEEAGESG----- 751
QY 491 QTHRAIFR----FVPRHED 505
DB 752 ---EGIFRDEFRVIRAED 767
```

```
RESULT 9
Q9HBL0
ID Q9HBL0 PRELIMINARY; PRT; 1735 AA.
AC Q9HBL0;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Tensin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RX MEDLINE=20480062; PubMed=11023826;
RA Chen H., Ishii A., Wong W.K., Chen L.B., Lo S.H.;
RT "Molecular characterization of human tensin."
RL Biochem. J. 351:403-411(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RA Lo S.H., Chen L.B.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF225896; AAG33700.1; -.
DR HSSP; P00519; IAB2.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR000050; PID_domain.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00462; PTB; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
DR PROSITE; PS00138; SUBTILASE_SER; UNKNOWN_1.
SQ SEQUENCE 1735 AA; 185677 MW; AC825375DD6CF87C CRC64;

Query Match 5.0%; Score 190.5; DB 4; Length 1735;
Best Local Similarity 21.2%; Pred. No. 4.5e-05;
Matches 185; Conservative 87; Mismatches 318; Indels 283; Gaps 39;

QY 12 GAASP-PAASPLGLHIASPNFR-----LTHDISLEEFDEDELSITDECGISLQCKD 64
DB 917 GYRSPVQCVSPALTIATLNPGRKPKPHLSYKAEFEMEGETSPSS----- 963
QY 65 TLSLRPPRAGLLSAGGGAGSLQAEMLQMD-----LIDATGDTPGAEDDEDDDE---- 115
DB 964 -----PPPSGVKSP-PGLAKTPLSALGLKPHNPADILLHPTGVTRRRITQPEDEGVVVR 1017
QY 116 -----ERAARRPGAGP-----PKAESGQEPASRGOGOGOGSGGDTYRPK- 159
DB 1018 LSEEPSYVESVARTAVAGPRAQDSEPKSFA--PATQAYGHEIPLRNGTLGGSFVSFSP 1075
QY 160 -----RPTTLNLP-----QVPRS-----QPTLNNSLGKKHMQDRVSRSS 197
DB 1076 LSTSSPILSADSTSVSGSPGESSDQGPRTPTQLLESGRSGSLGQPSAQRYQSS 1135
QY 198 PLKT-----GEQT-----PHEHICLSEELPP----- 219
DB 1136 PLPTVGVSSYSSPDYSLQHFSSSPSQAQAFVAGVHTVPGSQARHRTVGTNTPPSPGF 1195
QY 220 -----QSGPAPTTDRGTSTDSPCRSTATOMAP----- 247
DB 1196 GRRAINPSMAAPSPSLSHQMMGPPGPGFHGTVSSPQSSAATTPGSPSLCRHPAGVYQ 1255
QY 248 -----PGGP-----PAAPPGGR-----GHS 262
DB 1256 VSGLHNKVATTPGSPSLGRHPGAGHGNLGLHSLNATASPGSLGRHLGSGSVWPGSP 1315
```

QY 263 HRDR-IHYQADVRLATEEIVLTPVQRPDPAAPTSA--FLPPTESRMSVSSDPDPAAYP 319  
Db 1316 CLDRHVAIYGG-----YSTPEDRRPTLSROSSASGQAPSTPSFPVS-----PAYYP 1361  
QY 320 ---STAGRPHPSISEEBEGFDCILSSPERAPPGGGWRGSL-----GEPPPPRLASLSS 369  
Db 1362 GLSPSPATSPDSAAFRQSGTTPALPEKRMSVGDRAAGSLPNVATNGKVASPVPSGMS 1421  
QY 370 DTSALSVDVSKYTLVDEHAQLELVLRPCFGDYSDSDSATYVDCASVSSYESAIGE 429  
Db 1422 PSGG-STVSFSHTL-----PDFSKYMPDNPS---ETRAKYKFQDTSKYW 1463  
QY 430 EYEBAPROPACILSESDTDPEDVHFSGKFLNVFMGSRSSAESFGLFSCLINEEQ 489  
Db 1464 YKPEISRQAIALL-KOEPGAFIIRUSHSPRGAYGLAMKVSSPPPTI-----MQQNK 1517  
QY 490 EOTHRAIRFV---PRHEDELEVDLPLVLEQAEDYWEAYNMRTGARGVFPAYYAI 545  
Db 1518 DMTHLVHRFLIETGRGV-KLKGCPNEPNFGSLSALVYQHSIIPLALPCKLVIPNRDPT 1576  
QY 546 EVTKEPEHMAALAKNSWDQ---FRVKFLGSQVYPYHKGNDVLCAMQKIATTRLTVH 602  
Db 1577 DESKDSGPA---NSTADLKGQAACNVLFINSVDMESLTGPQ---AISK-ATSETLAAD 1629  
QY 603 FNPSSCV-LEISVRGVKIGVKAADDQAEKGNKCSHFPO-----LKNISFCGYHHPKNKY- 656  
Db 1630 PTPAATIVHFKVSAGGITL-----TONQRKLFRRHYPLNTVTFCDLDFQERKWM 1679  
QY 657 -----FGFIT-KHPADHRFACHVFVED 678  
Db 1680 KTEGGAPAKLFGFVARKOGSTTDNACHLFAELD 1712  
RESULT 10  
O95875  
ID O95875 PRELIMINARY; PRT; 2157 AA.  
AC O95875;  
DT 01-MAY-1999 (TremBLrel. 10, Created)  
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE BAT2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,  
RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,  
RA Lasky S., Hood L.;  
RT "Sequence of the human major histocompatibility complex class III  
region.";  
RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AF129756; AAD18086.1;  
DR EMBL; AF129756; AAD18086.1;  
SQ SEQUENCE 2157 AA; 228869 MW; 21B817F5B699B0DB CRC64;  
Query Match 5.0%; Score 190; DB 4; Length 2157;  
Best Local Similarity 22.7%; Pred. No. 6.6e-05;  
Matches 125; Conservative 34; Mismatches 180; Indels 212; Gaps 24;  
QY 3 ERES---GGLGGGAASPPAASP-----FLGLHIA-----SPNFRILTHDISLEEF 45  
Db 1263 ERNNAARGSEKPSLTLPASAPAEALTTVTIVAPAPRAAKSPDLNSQNSDQANEWE 1322  
QY 46 D-EDLSEITDECGISLQCKDTLSLRPPRAGLLSAGGGAGSRLOAEMQLMDLIDATGTP 104  
Db 1323 TASESSDFTSERRGDKAEPVLLTPKAVCTPGGGGAVPGISA-MSRGDLSQRAKDL 1381  
QY 105 GAEDDEDDDEERAARPCAGPKAESGOEPASRGOGSOGSGOGSDTYRKRPTTL 164  
Db 1382 KRFSFQRPMERQNRKRRPGGKAGSSGS--SSGGGGGPGGTGPRGD----- 1429  
QY 165 NLPFPVPRSQDTLNNNSLKGKHSWQDRVSRSSSPLKTGEQTPPHEHICLSEELPQSGPA 224

Db 1430 -----KRSWSPKNSRPP-----EERPP-----GLPLPP---PP 1456  
QY 225 PTTDRGTSTDSPCRSTA-----TQMAPPG-----GPEAAPGGRGHSHR 264  
Db 1457 PSSSAVFRIDQVIHNPAGIOQAOLSSROGSVTAPGGHPRKPGPPQAPQG----- 1509  
QY 265 DRIHQADVRLATEEIVLTPVQRPDPAAPTSAFLPPTESRMSVSSDP----- 313  
Db 1510 -----PSRPFTRYEQRV-----NSGLSSDPHFEEPGPMVRG 1542  
QY 314 -----DPAAY-----PSTAGRP--HPSISEEBEGFDCILSSPERAPPGGGWRGSLGEP 361  
Db 1543 VGGTPTSDSAGVSPFPKRRPRPKPELQEE-----SLP---PPHSSGFLGSKPEGPG 1593  
QY 362 PPRASLSDTSALS-----YDSVKYTLVDEHAQLELVSLRP-----CFGDY 403  
Db 1594 POAESRDTGTTEALTPHWNRL-----HTATSKRSTRPSSMEPMEPLSPFEDVAGTE 1645  
QY 404 SDESDSATVYDNCASVSS-----HTATSKRSTRPSSMEPMEPLSPFEDVAGTE 430  
Db 1646 MSQSDSGVDLSGDSQVSSGSCSORSSPDGGLKGAEEGPKRPGSSPLNAVPCGPPG-- 1703  
QY 431 YEAPRPQPPA 441  
Db 1704 -SEPPRRPPPA 1713  
RESULT 11  
Q96QC6  
ID Q96QC6 PRELIMINARY; PRT; 2157 AA.  
AC Q96QC6;  
DT 01-DEC-2001 (TremBLrel. 19, Created)  
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE BAT2 protein.  
GN BAT2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Shilina S., Tamiya G., Oka A., Inoko H.;  
RA Hiraoka M., Yamaguchi H., Imai K., Shimada J.;  
RP Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP000505; BAB63391.1;  
SQ SEQUENCE 2157 AA; 228858 MW; C500E03BE15382C9 CRC64;  
Query Match 5.0%; Score 190; DB 4; Length 2157;  
Best Local Similarity 22.7%; Pred. No. 6.6e-05;  
Matches 125; Conservative 34; Mismatches 180; Indels 212; Gaps 24;  
QY 3 ERES---GGLGGGAASPPAASP-----FLGLHIA-----SPNFRILTHDISLEEF 45  
Db 1263 ERNNAARGSEKPSLTLPASAPAEALTTVTIVAPAPRAAKSPDLNSQNSDQANEWE 1322  
QY 46 D-EDLSEITDECGISLQCKDTLSLRPPRAGLLSAGGGAGSRLOAEMQLMDLIDATGTP 104  
Db 1323 TASESSDFTSERRGDKAEPVLLTPKAVCTPGGGGAVPGISA-MSRGDLSQRAKDL 1381  
QY 105 GAEDDEDDDEERAARPCAGPKAESGOEPASRGOGSOGSGOGSDTYRKRPTTL 164  
Db 1382 KRFSFQRPMERQNRKRRPGGKAGSSGS--SSGGGGGPGGTGPRGD----- 1429  
QY 165 NLPFPVPRSQDTLNNNSLKGKHSWQDRVSRSSSPLKTGEQTPPHEHICLSEELPQSGPA 224  
Db 1430 -----KRSWSPKNSRPP-----EERPP-----GLPLPP---PP 1456









GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2002, 08:48:50 : Search time 18 Seconds

(without alignments)  
1162.205 Million cell updates/sec

Title: US-09-966-561-2

Perfect score: 3792

Sequence: 1 MAERESGLGGGAASPPAAS.....QQFYKQFVEVTCPTDIYLE 711

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2.6/ptodata/1/iaa/5A.COMB.pep.\*
- 2: /cgn2.6/ptodata/1/iaa/5B.COMB.pep.\*
- 3: /cgn2.6/ptodata/1/iaa/6A.COMB.pep.\*
- 4: /cgn2.6/ptodata/1/iaa/6B.COMB.pep.\*
- 5: /cgn2.6/ptodata/1/iaa/PTCUS.COMB.pep.\*
- 6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3503	92.4	714	2 US-08-859-201-2	Sequence 2, Appli
2	3263.5	86.1	660	3 US-08-819-177-1	Sequence 1, Appli
3	3245	85.6	659	4 US-09-562-737-11	Sequence 11, Appl
4	2875	75.8	659	4 US-09-562-737-20	Sequence 20, Appl
5	2864	75.5	659	4 US-09-562-737-16	Sequence 16, Appl
6	2846	75.1	659	4 US-09-562-737-13	Sequence 13, Appl
7	2845	75.0	659	4 US-09-562-737-17	Sequence 17, Appl
8	2829	74.6	659	4 US-09-562-737-12	Sequence 12, Appl
9	2824	74.5	659	4 US-09-562-737-14	Sequence 14, Appl
10	2817	74.3	659	4 US-09-562-737-15	Sequence 15, Appl
11	2816	74.3	659	4 US-09-562-737-18	Sequence 18, Appl
12	2814	74.2	659	4 US-09-562-737-19	Sequence 19, Appl
13	1655	43.6	314	2 US-08-859-201-8	Sequence 8, Appli
14	1427	37.6	270	2 US-08-859-201-4	Sequence 4, Appli
15	1279.5	33.7	830	4 US-09-562-737-31	Sequence 31, Appl
16	1116.5	29.4	830	4 US-09-562-737-36	Sequence 36, Appl
17	1111.5	29.3	830	4 US-09-562-737-33	Sequence 33, Appl
18	1103.5	29.1	829	4 US-09-562-737-40	Sequence 40, Appl
19	1097.5	28.9	830	4 US-09-562-737-38	Sequence 38, Appl
20	1093.5	28.8	830	4 US-09-562-737-34	Sequence 34, Appl
21	1083.5	28.6	830	4 US-09-562-737-32	Sequence 32, Appl
22	1075.5	28.4	830	4 US-09-562-737-37	Sequence 37, Appl
23	1074	28.3	830	4 US-09-562-737-39	Sequence 39, Appl
24	1060.5	28.0	830	4 US-09-562-737-35	Sequence 35, Appl
25	751	19.8	155	3 US-08-819-177-4	Sequence 4, Appli
26	392	10.3	79	3 US-08-819-177-16	Sequence 16, Appl
27	389	10.3	77	3 US-08-819-177-17	Sequence 17, Appl

28 359 9.5 76 3 US-08-819-177-15 Sequence 15, Appli  
29 327 8.6 68 3 US-08-819-177-18 Sequence 18, Appli  
30 305 8.0 59 3 US-08-819-177-19 Sequence 19, Appli  
31 261 6.9 49 3 US-08-819-177-20 Sequence 20, Appli  
32 260 6.9 50 3 US-08-819-177-13 Sequence 13, Appli  
33 207 5.5 39 3 US-08-819-177-21 Sequence 21, Appli  
34 186 4.9 696 3 US-08-906-865-4 Sequence 4, Appli  
35 186 4.6 696 4 US-09-129-668-4 Sequence 2, Appli  
36 175.5 4.6 1243 2 US-08-557-139-2 Sequence 2, Appli  
37 169.5 4.5 1185 4 US-09-041-886-23 Sequence 23, Appli  
38 169 4.5 1234 2 US-08-317-310A-15 Sequence 15, Appli  
39 169 4.5 1234 5 PCT-US95-13041-15 Sequence 15, Appli  
40 167.5 4.4 1958 1 US-07-945-283-2 Sequence 2, Appli  
41 166.5 4.4 1155 1 US-08-094-948A-29 Sequence 29, Appli  
42 166.5 4.4 1155 5 PCT-US96-09319-29 Sequence 29, Appli  
43 165.5 4.4 548 2 US-08-469-412A-2 Sequence 2, Appli  
44 165.5 4.4 548 4 US-09-021-715-2 Sequence 2, Appli  
45 165 4.4 1291 4 US-09-150-460B-10 Sequence 10, Appli

#### ALIGNMENTS

RESULT 1  
US-08-859-201-2  
; Sequence 2, Application US/08859201  
; Patent No. 5880261  
; GENERAL INFORMATION:  
; APPLICANT: Waerber, G  
; APPLICANT: Nicod, P  
; TITLE OF INVENTION: Transcription Factor Islet-Brain 1 (IB1)  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Luann Cserf  
; STREET: 750 Arlmo Avenue  
; CITY: Oakland  
; STATE: California  
; COUNTRY: USA  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION NUMBER: US/08/859,201  
; FILING DATE: 20-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9706731.8  
; FILING DATE: 03-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9709920.4  
; FILING DATE: 15-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Luann Cserf  
; REGISTRATION NUMBER: 31,822  
; REFERENCE/DOCKET NUMBER: ME A9702  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 714 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-859-201-2

Query Match 92.4%; Score 3503; DB 2; Length 714;  
Best Local Similarity 93.3%; Pred. No. 4.9e-294;  
Matches 664; Conservative 11; Mismatches 31; Indels 6; Gaps 5;

QY 1 MAERESGLGGGAASPPAASPPFLGLHIAASPPNFRHLTHDLSLEEFEDDLSEITDCGSL 60  
||||| || ||||||||| ||||||||| ||||||||| ||||||||| |||||||||  
Db 8 MAERES-GLSGGAASPPAASPPFLGLHIAASPPNFRHLTHDLSLEEFEDDLSEITDCGSL 66  
QY 61 QCKDTLSLRPPRAGLLSAG-GGAGSRLQAEMQLQNDLIDATGDTGCAEDDEDDDEERAA 119

```

; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07917/037001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; FEATURE:
; OTHER INFORMATION: JIP-1 protein
US-08-819-177-1

```

```

Query Match      86.1%; Score 3263.5; DB 3; Length 660;
Best Local Similarity 87.2%; Pred. No. 2.1e-273;
Matches 621; Conservative 12; Mismatches 26; Indels 53; Gaps 4;

```

```

QY 1 MAERESGLGGGAASPPAASPFLGLHTIASPPNFRLLHDISLEEFDEDLSEITDDECGISL 60
DB 1 MAERES-GLGGGAASPPAASPFLGLHTIASPPNFRLLHDISLEEFDEDLSEITDDECGISL 59
QY 61 QCKDTLSLRPPRAGLLSAG-GGGAGSRLQAEMLQMDLIDATGTPGAEDDEDEDEERA 119
DB 60 QCKDTLSLRPPRAGLLSAGSRLQAEMLQMDLIDAAAGTGAEDDEDEDEELAA 119
QY 120 RPPGAPPKAESGGEAPASRGOGSGOGSGDGYRKRPTTLNLFQVPRSDTLNN 179
DB 120 QPPGVGPPKAESNODPAPR---SQGGPGTGSDDYRKRPTTLNLFQVPRSDTLNN 175
QY 180 NSLKKHSHQDVRSSRSPLKTEGQTPPHHICLSBELPPQSGPAPTDRGTSTDSPCR 239
DB 176 NSLKKHSHQDVRSSRSPLKTEGQTPPHHICLSBELPPQSGPAPTDRGTSTDSPCR 235
QY 240 STATQMAPPGGPPAAPPGGRGSHRDRHYQADVRLATEEETIYLTVPQRPDAAPTSAF 299
DB 236 SAATQMAPPGGPPATAPGGRGSHRDRHYQADVRLATEEETIYLTVPQRPDAAPTSTF 295
QY 300 LPPTESRMSVSSDDPDAAYSTAGRPHPSISEEFGFDCLSSPERAEPGGGWRGSLGEP 359
DB 296 MPPTESRMSVSSDDPDAAYSVTAGRPHPSISEEFGFDCLSSPERAEPGGGWRGSLGEP 355
QY 360 PPPRASLSDDTSALSYDSVKYTLVVDHAQLVSLRPFCDYSDSDSATVYDNCASV 419
DB 356 PPPRASLSDDTSALSYDSVKYTLVVDHAQLVSLRPFCDYSDSDSATVYDNCASA 415
QY 420 SSPYESAIGEEYEAPRPPACILSDSTDPEDPVHFSKFLNFMVMSGRSSSAESFGL 479
DB 416 SSPYESAIGEEYEAPRPPACILSDSTDPEDPVHFSKFLNFMVMSGRSSSAESFGL 475
QY 480 FSCIINGEEQETHRAIFRFVPRHEDELEVDPLLLVQLQAEDEYWEAYNNMRTGARGVF 539
DB 476 FSCVINGEEQETHRAIFRFVPRHEDELEVDPLLLVQLQAEDEYWEAYNNMRTGARGVF 535
QY 540 PAYTAIEVTKEPEHMAALAKNSDWDQFRVYKFLGSLVQVPHKGNVLCAMOKIATTRL 599
DB 536 PAYTAIEVTKEPEHMAALAKN----- 556
QY 600 TVHFNPPSSCVLEISVRGVKIGVADDSQBAKGNKCSHFFQLKNISFCGYHPKNNKYEGF 659
DB 557 -----SCVLEISVRGVKIGVADDALEAKGNKCSHFFQLKNISFCGYHPKNNKYEGF 608
QY 660 ITKHPADHRFACHVFVSEDSTKALAESVGRAFOQFYKQFVEYTCPTEDIYLE 711
DB 609 ITKHPADHRFACHVFVSEDSTKALAESVGRAFOQFYKQFVEYTCPTEDIYLE 660

```

```

RESULT 3
US-09-562-737-11
; Sequence 11, Application US/09562737

```

```

RESULT 2
US-08-819-177-1
; Sequence 1, Application US/08819177
; Patent No. 6043083
; GENERAL INFORMATION:
; APPLICANT: Davis, Roger J.
; APPLICANT: Dickens, Martin
; TITLE OF INVENTION: INHIBITORS OF THE JNK SIGNAL
; TITLE OF INVENTION: TRANSDUCTION PATHWAY AND METHODS OF USE
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,177
; FILING DATE: 28 April 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```



; Patent No. 6428967  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gotthardt, Michael  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/09/562.737  
; CURRENT FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 659  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Sequence  
US-09-562-737-11

Query Match 85.6%; Score 3245; DB 4; Length 659;  
Best Local Similarity 87.1%; Pred. No. 8.4e-272;  
Matches 620; Conservative 12; Mismatches 26; Indels 54; Gaps 5;

QY	1	MAERESGGLGGGAAPPAAAPFLGLHIAAPPNFRLTHTDISLEEFEDDLSEITDECCISL	60
DB	1	MAERES-GLGGGAAPPAAAPFLGLHIAAPPNFRLTHTDISLEEFEDDLSEITDECCISL	59
QY	61	QCKDTLSLRPRRAGLLSAG-GGAGSRLQAEMLQMDLIDATGPTGAEDDEEDDEERAA	119
DB	60	QCKDTLSLRPRRAGLLSAGSGSRLQAEMLQMDLIDAAAGDTPGAEDDEEDDEELAA	119
QY	120	RRGAGPPKAESQBPASRGQSGOGSGDTPYRKPRTTLNLFPPVPRSQDTLNN	179
DB	120	RRGAGPPKAESQBPASRGQSGOGSGDTPYRKPRTTLNLFPPVPRSQDTLNN	175
QY	180	NSLGGKHSQDVRSSSPLKTGEOTPPHEHICLSEELPPQSGPAPITDRTGSTDSPCRR	239
DB	176	NSLGGKHSQDVRSSSPLKTGEOTPPHEHICLSDLPQSGVPVPTQDRTGSTDSPCRR	235
QY	240	STATQMAPGGPPAAPPGGRGSHDRTHYQADVRLATEEIIYTPVQRPDAEPTSAF	299
DB	236	SAATQMAPGGPPAAPPGGRGSHDRTHYQADVRLATEEIIYTPVQRPDAEPTSTF	295
QY	300	LPPTSRMSVSDPDPAAYPTAGRPHPHSISEEEGFCCLSSPERAPPGGGWRGSLGEP	359
DB	296	MPPTSRMSVSDPDPAAYPTAGRPHPHSISEEEDGFCCLSSPERAPPGGGWRGSLGEP	355
QY	360	PPPRASLSSTDSALSYDVKYTLVVDHQAOLVSLRCPGCDYSDSDSATVYDNCASV	419
DB	356	PPPRASLSSTDSALSYDVKYTLVVDHQAOLVSLRCPGCDYSDSDSATVYDNCASA	415
QY	420	SSPYSAIGEEYEAPRPPQACLSDESTDPDPVHFSKFLNFMGSRSSSAESFGL	479
DB	416	SSPYSAIGEEYEAPRPPQACLSDESTDPDPV-FSKFLNFMGSRSSSAESFGL	474
QY	480	FSCIINGEQEOTHRAIRFRVPRHDELELEVDPLVLAQAEYWEAYNMRGARGVF	539
DB	475	FSCIINGEQEOTHRAIRFRVPRHDELELEVDPLVLAQAEYWEAYNMRGARGVF	534
QY	540	PAYIAEVTKEPEHMAALAKNSDWQDFRVKFLGVSQVYPYHKGNDVLCAMQKIATTTRL	599
DB	535	PAYIAEVTKEPEHMAALAKN-----	555
QY	600	TVHFNPSSCVLEISVRGVKIGVKADDSQAEKGNKCSHFFOLKNISFCGYHPKNNKYFGF	659
DB	556	-----SCVLEISVRGVKIGVKADDALEAKGNKCSHFFOLKNISFCGYHPKNNKYFGF	607
QY	660	ITKHPADHRFACHVFSVSDSTKALAESVGRAPQOQFYKOFVEYTCPTEDIYLE	711
DB	608	ITKHPADHRFACHVFSVSDSTKALAESVGRAPQOQFYKOFVEYTCPTEDIYLE	659

RESULT 4

US-09-562-737-20  
; Sequence 20, Application US/09562737  
; Patent No. 6428967  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gotthardt, Michael  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/09/562.737  
; CURRENT FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 659  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Sequence  
US-09-562-737-20

Query Match 75.8%; Score 2875; DB 4; Length 659;  
Best Local Similarity 78.8%; Pred. No. 7.6e-240;  
Matches 561; Conservative 18; Mismatches 79; Indels 54; Gaps 5;

QY	1	MAERESGGLGGGAAPPAAAPFLGLHIAAPPNFRLTHTDISLEEFEDDLSEITDECCISL	60
DB	1	MAERES-GLGGAAPPAAAPFLGLHIAAPPNFRLTHTDISLEEFEDDLSEITDECCISL	59
QY	61	QCKDTLSLRPRRAGLLSAG-GGAGSRLQAEMLQMDLIDATGPTGAEDDEEDDEERAA	119
DB	60	QCKDTLSLRPRRAGLLSAGSGSRLQAEMLQMDLIDAGDTPGAEDDEEDDEELAA	119
QY	120	RRGAGPPKAESQBPASRGQSGOGSGDTPYRKPRTTLNLFPPVPRSQDTLNN	179
DB	120	RRGAGPPKAESQBPASRGQSGOGSGDTPYRKPRTTLNLFPPVPRSQDTLNN	175
QY	180	NSLGGKHSQDVRSSSPLKTGEOTPPHEHICLSEELPPQSGPAPITDRTGSTDSPCRR	239
DB	176	NSLGGKHSQDVRSSSPLKTGEOTPPHEHICLFDLPQSGPQDRTGSTDSPCRR	235
QY	240	STATQMAPGGPPAAPPGGRGSHDRTHYQADVRLATEEIIYTPVQRPDAEPTSAF	299
DB	236	SAATQMAPGGPPAAPPGGRGSHDRTHYQADVRLATEEIIYTPVQRPDAEPTSTF	295
QY	300	LPPTSRMSVSDPDPAAYPTAGRPHPHSISEEEGFCCLSSPERAPPGGGWRGSLGEP	359
DB	296	MPPTSRMSVSDPDPAAYPTAGRPHPHSISEEDVGFCLSSPERAPPGGGWRGSLGEP	355
QY	360	PPPRASLSSTDSALSYDVKYTLVVDHQAOLVSLRCPGCDYSDSDSATVYDNCASV	419
DB	356	PPPRASLSSTDSALSYDVKYTLVVDHQAOLVSLRCPGCDYSDSDSATVYHNCASA	415
QY	420	SSPYSAIGEEYEAPRPPQACLSDESTDPDPVHFSKFLNFMGSRSSSAESFGL	479
DB	416	SSPYSAIGEEYEAPRPPQACLSDESTDPDPV-FSKFLNFMGSRSSSAESFGL	474
QY	480	FSCIINGEQEOTHRAIRFRVPRHDELELEVDPLVLAQAEYWEAYNMRGARGVF	539
DB	475	FSCIINGEQEOTHRAIRFRVPRHDELELEVDPLVLAQAEYWEAYNMRGARGVF	534
QY	540	PAYIAEVTKEPEHMAALAKNSDWQDFRVKFLGVSQVYPYHKGNDVLCAMQKIATTTRL	599
DB	535	PAYIAEVTKEPEHMAALAKN-----	555
QY	600	TVHFNPSSCVLEISVRGVKIGVKADDSQAEKGNKCSHFFOLKNISFCGYHPKNNKYFGF	659
DB	556	-----SCVLEISVRGVKIGVKADDALEAKGNKCSHFFOLKNISFCGYHPKNNKYFGF	607
QY	660	ITKHPADHRFACHVFSVSDSTKALAESVGRAPQOQFYKOFVEYTCPTEDIYLE	711
DB	608	ITKHPADHRFACHVFSVSDSTKALAESVGRAPQOQFYKOFVEYTCPTEDIYLE	659





QY 600 TVHFNPPSCVLEISVRGKIGVKADDSQEAQKGNKSHFFQFKNLSFCGYHPKNNKYFGF 659  
Db 556 -----SCVLEILVRGKIGVKMDDALEAKGNKNSHFFQFKNLSFCGYHPKNNKYFGF 607  
QY 660 ITKHPADHRFACHVFVSEDSKALAESVGRFAQFOQFKQFVEYTCPTEDIYLE 711  
Db 608 ITKHSADHRFACHVTVSEDSKALVESVGRFAQFOQFKQFVEYTCYCTEDIYLE 659

RESULT 9  
US-09-562-737-14  
; Sequence 14, Application US/09562737  
; Patent No. 6428967  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/09/562,737  
; CURRENT FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 659  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Sequence  
US-09-562-737-14

Query Match 74.5%; Score 2824; DB 4; Length 659;  
Best Local Similarity 78.7%; Pred. No. 1.9e-235;  
Matches 560; Conservative 20; Mismatches 78; Indels 54; Gaps 5;

QY 1 MAERESGLGGGAASPPAASPFLGLHIAAPPNFRFLTHDISLEEFDEDLSEITDECGISL 60  
Db 1 MAEAS-GLGGAANDPPAASPFLGELIASPPNFRFLTHDISLEEFDEDLSEITDHGGISL 59  
QY 61 QCKDITLSRPPRAGLLSAGGAG-SRLQAEMLQMLDIDATGDTPGAEDDEEDDEERAA 119  
Db 60 QCKDITLSRPPRAGKLSAGSSGALSRLQAEMLQMLDIDAAAGTGAEDDEEERELAA 119  
QY 120 RPPGAPPKAESQEPASRGQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQ 179  
Db 120 QRPQSGPPKAEASNQTPAPR----SQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQ 175  
QY 180 NSLGGKHSQDRVSRSSSPKLTGEOTPPHEHICLSEELPPQSGPAPTTDRGTSTDSPCR 239  
Db 176 NSLGGKHSQDRVSRSSSELKLTGEOTPPHEHICLSEELPPQSGPAPTTQHRGTSTDSPCR 235  
QY 240 STATQMAPPGPPAAPPGGGRGSHRDRIHYQADVRLAEATEEILYLPVQRPDPAEPTSAF 299  
Db 236 SAATQMAPKSGPPATAPGURGSHRDRIHYQADVRLAEATEEILYLPVQRPDPAEPTSRF 295  
QY 300 LPPTESRMSVSSDDPPAAYSTAGRPHPSISEEEGFDCLSSPERAEPPGGGWRGSLGRP 359  
Db 296 MPPTESRMTVSSDDPPAASVTAAGRPHPYISEEDGFDALSSPERAEPPGGGWRGSLGEP 355  
QY 360 PPPPRASLSDTSALSVDYKTLVVDHAQLELVSLRCPGQDYSDESASVYONCASV 419  
Db 356 PPPPRASLSDTSALSVDYKTLVVDHAQLELVSLRCKFCGQDYSDESASVYONCAMA 415  
QY 420 SSPYESAIGEEYEAAPRPPACLSDESTPDPVHFSSKFLNVFMGSRSSSAESFGL 479  
Db 416 SSPYESAINEEYEAAPQPPCTCLSEDSRDPEDV-FSKSFLNVFMGSRSSSAESFGL 474  
QY 480 FSCIINGEQEOTHRAIFRVPRHEDELELVDDPLLVELQAEADYWEAYNMTGARGVF 539  
Db 475 FSCVINGEWEQTHRAIFRVPRHEDELEAEVDDPLLVEDQAEADYWEAYNMTGARGVG 534  
QY 540 PAYYAIETVKEPEHMAALAKNSDWQDFRVKFLGSQVQPYHKGNDVLCAAMOKIATTRL 599  
|||||

Db 535 PAYYAIETVTHEPEHMAALAIN----- 555  
QY 600 TVHFNPPSCVLEISVRGKIGVKADDSQEAQKGNKSHFFQFKNLSFCGYHPKNNKYFGF 659  
Db 556 -----SCVLEISVRGKIGVKADLALEAKGNKCMHFFQFKNLSFCGYHPKNNKYFGF 607  
QY 660 ITKHPADHRFACHVFVSEDSKALAESVGRFAQFOQFKQFVEYTCPTEDIYLE 711  
Db 608 ITKHPARHRFACHVFVTEDSKALAEVVGRAFOQFVQFVEYTCPTEDIYLE 659

RESULT 10  
US-09-562-737-15  
; Sequence 15, Application US/09562737  
; Patent No. 6428967  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/09/562,737  
; CURRENT FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 659  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Sequence  
US-09-562-737-15

Query Match 74.3%; Score 2817; DB 4; Length 659;  
Best Local Similarity 78.4%; Pred. No. 7.7e-235;  
Matches 560; Conservative 18; Mismatches 78; Indels 58; Gaps 6;

QY 1 MAERESGLGGGAASPPAASPFLGLHIAAPPNFRFLTHDISLEEFDEDLSEITDECGISL 60  
Db 1 MAERAS-GLGGAANDPPAASPFLGELIASPPNFRFLTHDISLEEFDEDLSEITDHGGISL 59  
QY 61 QCKDITLSRPPRAGLLSAG-GGGAGSRLQAEMLQMLDIDATGDTPGAEDDEEDDEERAA 119  
Db 60 QCKDITLSRPPRAGKLSAGSSGALSRLQAEMLQMLDIDAAAGTGAEDDEEEDQELAA 119  
QY 120 RPPGAPPKAESQEPASRGQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQ 177  
Db 120 QRPQSGPPKAEASNQTPAPR----SQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQ 173  
QY 178 NNLSLGGKHSQDRVSRSSSPKLTGEOTPPHEHICLSEELPPQSGPAPTTDRGTSTDSPC 237  
Db 174 NYNSLGGKHSQDRVSRSSSPKLTGEOTPPHEHICLSEELPPQSGPAPTTDRGTSTDSPC 233  
QY 238 RRSTATQMAPPGPPAAPPGGGRGSHRDRIHYQADVRLAEATEEILYLPVQRPDPAEPTS 297  
Db 234 RISAATQMAPKSGPPATAPGGLGSHRDRIHQADVRLAEATEEILYLPVQRPDPAEPTS 293  
QY 298 AFLPPTESRMSVSSDDPPAAYSTAGRPHPSISEEEGFDCLSSPERAEPPGGGWRGSLG 357  
Db 294 TRMPTESRMTSSSDPPAAYTVTAGRPHPSYSEDEGFCDCSWSPERAEPPYGGWRGSLG 353  
QY 358 EPPPPRASLSDTSALSVDYKTLVVDHAQLELVSLRCPGQDYSDESASVYDNCA 417  
Db 354 EAPPPPRASLSDTSALSVDYKTLVVDHAQLELVSLRCPGQDYSDESASVYDNCA 413  
QY 418 SVSSPYESAIGEEYEAAPRPPACLSDESTPDPVHFSSKFLNVFMGSRSSSAESF 477  
Db 414 SSSPYESAIGEEYEAAPQRLPTCLSEDSRDPEDV-FSKKLNLFVFMGSRSSSAESF 472  
QY 478 GLFSCIINGEQEOTHRAIFRVPRHEDELELVDDPLLVELQAEADYWEAYNMTGARG 537  
Db 473 GLRSCVINGEWEQTHRAIFRVPRHEDELELVDDPLLVELWAEDYWEAYNMTGARG 532  
|||||

QY 538 VPPAYATVETKEPEHMAALAKNSDWQDFRVKFLGSGVQVPHYHKGNDVLCAMOKIATTR 597  
Db 533 VPAAYATVETKDPHEMAALAK-----554  
QY 598 RLVHFNPPSCVLSEISVGVKIGVKADDSQEAAGKNCSEHFFQKLNISFCGYHPKNNKYF 657  
Db 555 -----ESCVLSEISRVFKIGVKADDDGLEAKGKCSIFFQKLNISFKGYHPKNNKYL 605  
QY 658 GFITKHPADHRFACHVFSVSDSTALAESVGRFQFQFYKQFVETCTPTEDIYLE 711  
Db 606 GFITKHPADMRFACHVFSVSDSTALAESQGRFQFQFYKREVEYTCPTESIYLE 659  
RESULT 11  
US-09-562-737-18  
; Sequence 18, Application US/09562737  
; Patent No. 6428967  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/09/562,737  
; CURRENT FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 659  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Sequence  
US-09-562-737-18  
Query Match 74.3%; Score 2816; DB 4; Length 659;  
Best Local Similarity 78.2%; Pred. No. 9.4e-235;  
Matches 557; Conservative 18; Mismatches 83; Indels 54; Gaps 5;  
QY 1 MAERESGLGGGAASPPAASPFGLGHIAFPNFRLTHDISLEEFEDDLSEITDECGISL 60  
Db 1 MAERESGA-GGGAASPPADSPFLGHIAFPNFRLTHDFSLEEFEDDSEITDECGIHL 59  
QY 61 QCKDTLSLRPRAGLLSA-GGGAGSRLQAEMLQMDLLDADTDPGAEDEDEDEERAA 119  
Db 60 QCKDTLSLRPRAGLLSAKSGSAGSRLQAEMLQMDLLMAAGDTPGAENDEEDEDLQA 119  
QY 120 RRPAGPPKAESQBPASRGQSGQSGQSGDTPYRKRPPTTLNLPQVPRSQDTLNN 179  
Db 120 QRPVGPVPAESNQDPAP-----SSQGGPGTGTGTYPYRKRPVTLNLPQVPRSQDTLNN 175  
QY 180 NSLKKHSHWQDRVSRSSPLKGTGEQTPPHEHICLSEELPPQSGAPPTDRTSTDSPCRR 239  
Db 176 NSYGGKKHSHWQDRASRSSLKGTDEQTPPHEHIELSDELPPQGFVPTQDRGTDSPCRR 235  
QY 240 STATMAPGGPPAAPGGGRGSHDRHYQADVRLATEEILYTPVQRPDAAEPTSAF 299  
Db 236 SAHTQMAPGGPIATAPGGRGHKHRRHYQALVRLATEEILMTLPVQRPDAAEPTSTF 295  
QY 300 LPPTESRMVSSDDPPAAYSTAGRPHPSISEEERGFCCLSPERAEPGGGWRGSLGEP 359  
Db 296 MPQTESRMVSSRDPDPAAYSVTSGRPHPSISETDSEDFCLSSVERAEPGGGWRGSLGEP 355  
QY 360 PPPRASLSSDTSALSYDSVKYTLVVDDEHAQLELVSLRCPGDIYSDSDSATVYDNCASV 419  
Db 356 PPAPRASLSSDTSALSYDSVKYELVVDDEHAQLFLVSLRCPGFGYSDESSTATHYDNCASA 415  
QY 420 SSPYESAIGEEYEAPRQPPACLSDESTPDEPDVHFSKFLNFMVSGRGRSSSAESFGL 479  
Db 416 SSIIYESAIGEEYEAPRQPPPLLEDSTPDMDV-FSKKFLNNFMVSGRGRSSSAESFGL 474  
QY 480 FSCIINGEQEQTHRAIFRVPRHDELELEVDVPLLVLAQEWYWEAYNNMRTGARGVF 539  
Db 411 FSCIINGEQEQTHRAIFRVPRHDELELEVDVPLLVLAQEWYWEAYNNMRTGARGVF 539

Db 475 FSCRINGEEHQTSRAIFRVPRHDELELEVDVPLLVLAQEWYWEAYNNMRTGARGVF 534  
QY 540 PAYATVETKEPEHMAALAKNSDWQDFRVKFLGSGVQVPHYHKGNDVLCAMOKIATTRRL 599  
Db 535 PAYAAVETKEPEDMAALAK-----555  
QY 600 TVHFNPPSCVLSEISVGVKIGVKADDSQEAAGKNCSEHFFQKLNISFCGYHPKNNKYF 659  
Db 556 -----SCELVSEISRVFKIGVKADDDGLEAKGKCSIFFQKLNISFKGYHPKNNKYG 607  
QY 660 ITKHPADHRFACHVFSVSDSTALAESVGRFQFQFYKQFVETCTPTEDIYLE 711  
Db 608 KTKHPADHRFLCHVFSVSDSMKALAESVGRNFOQFQFYKQFVQVTCPTEDIYRE 659  
RESULT 12  
US-09-562-737-19  
; Sequence 19, Application US/09562737  
; Patent No. 6428967  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/09/562,737  
; CURRENT FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 659  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Sequence  
US-09-562-737-19  
Query Match 74.2%; Score 2814; DB 4; Length 659;  
Best Local Similarity 78.2%; Pred. No. 1.4e-234;  
Matches 556; Conservative 18; Mismatches 83; Indels 54; Gaps 5;  
QY 1 MAERESGLGGGAASPPAASPFGLGHIAFPNFRLTHDISLEEFEDDLSEITDECGISL 60  
Db 1 MAERES-GLAGGAASPPADPFGLGHIAFPNFRLTHDIFLEEFEDDLSEITDECGISH 59  
QY 61 QCKDTLSLRPRAGLLSAG-GGGAGSRLQAEMLQMDLLDADTDPGAEDEDEDEERAA 119  
Db 60 QCKDTLSLRPRAGLLSAGSAGSRLQAEMLQMDLLDADTDPGAEDEDEDEDELAQ 119  
QY 120 RRPAGPPKAESQBPASRGQSGQSGQSGDTPYRKRPPTTLNLPQVPRSQDTLNN 179  
Db 120 QRPVGPVPAESNQDPAP-----TQGGPGTGTGTYPYRKRPVTLNLPQVPRSQDTLNN 175  
QY 180 NSLKKHSHWQDRVSRSSPLKGTGEQTPPHEHICLSEELPPQSGAPPTDRTSTDSPCRR 239  
Db 176 NSLAKHSHWQDRVSRSSPLKGTGEQTPPHEHICLSEELPPQSGSHVPTQDRGTSDSPCRR 235  
QY 240 STATMAPGGPPAAPGGGRGSHDRHYQADVRLATEEILYTPVQRPDAAEPTSAF 299  
Db 236 SAAKQMAPPPGCPPLTAPGGRGHSMRDRHYQADVRLATEEILYTPVQRPDREPTSTF 295  
QY 300 LPPTESRMVSSDDPPAAYSTAGRPHPSISEEERGFCCLSPERAEPGGGWRGSLGEP 359  
Db 296 MPPEESRMVSSDDPPAAYSVTAVRPHPSISEEWEFGDCLSSPERAEPGGGWRGSLGEP 355  
QY 360 PPPRASLSSDTSALSYDSVKYTLVVDDEHAQLELVSLRCPGDIYSDSDSATVYDNCASV 419  
Db 356 PPDRASLSSDTSALSYDSVKYTFVVDDEHAQLELVSLRCPGDIYSDSDSATVYDNCASA 415  
QY 420 SSPYESAIGEEYEAPRQPPACLSDESTPDEPDVHFSKFLNFMVSGRGRSSSAESFGL 479  
Db 416 SSPKESAIGEEYELAPQRPPTCMSESTPDENP-FSKKFLNFMVSGRGRSSSAESFGL 474

```
QY 480 FSCIINGEQETHRAIRFRVPRHDELELEVELEVDPLLVLAEDYWEAYNMRTGARGVF 539
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 475 FSCVNGEHEQETHRAIRFRVPRHVDLELEVELEVDMLLVLAEDAWYAYNMRTDARGVF 534

QY 540 PAYIAEVTKEPEHMAALAKNSDWVQDFRVKFLGSVQVPYHKGNDVLCAMQKIATTTT 599
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 535 PAYIELEVTKEPEHFAALAKN----- 555

QY 600 TVHFNPSSCVLEISVRGVKIGVADDSQEAAGKNCSHFFOLKNISFCGYHPKNNKYFGF 659
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 556 -----SCVGEISVRGVKIHVKADDALEALGNKCSHFFOLKNISFCGYHLKNNKYFGF 607

QY 660 ITKHPADHRFACHVSEDSKALAESVGRAPQOQFYKQFVETCTPTEDIYL 710
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 608 IMKHPADHRFANHVEFVSEDSQALAESVGRARQOQFYKQFVETCTPTEDIYL 658

RESULT 13
US-08-859-201-8
; Sequence 8, Application US/08859201
; Patent No. 5880261
; GENERAL INFORMATION:
; APPLICANT: Waeber, G
; APPLICANT: Nicod, P
; TITLE OF INVENTION: Transcription Factor Islet-Brain 1 (IB1)
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Luann Cserr
; STREET: 750 Arimo Avenue
; CITY: Oakland
; STATE: California
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,201
; FILING DATE: 20-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9706731.8
; FILING DATE: 03-APR-1997
; APPLICATION NUMBER: GB 9709920.4
; FILING DATE: 15-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Luann Cserr
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: ME A9702
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 314 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-859-201-8
Query Match 43.6%; Score 1655; DB 2; Length 314;
Best Local Similarity 97.1%; Pred. No. 5.8e-135;
Matches 305; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 398 PCFGYSDSDSATVYDNCASVSPYSAIGEYEEAPRPPOACLSDSTDPDPVHFS 457
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 PCFGYSDSDSATVYDNCASVSPYSAIGEYEEAPRPPTCLSDSTDPDPVHFS 60

QY 458 KKFLNVMSGRRSSAESFGLFSCIINGEQETHRAIRFRVPRHDELELEVDPLLV 517
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 KKFLNVMSGRRSSAESFGLFSCVINGEQETHRAIRFRVPRHDELELEVDPLLV 120

QY 518 ELQADYWEAYNMRTGARGVFPAYIAEVTKEPEHMAALAKNSDWVQDFRVKFLGSVQV 577
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 ELQADYWEAYNMRTGARGVFPAYIAEVTKEPEHMAALAKNSDWIDQDFRVKFLGSVQV 180
```

```
QY 578 PYHKGNDVLCAMQKIATTTTTRRLTVHFNPSSCVLEISVRGVKIGVADDSQEAAGKNCSH 637
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 PYHKGNDVLCAMQKIATTTTTRRLTVHFNPSSCVLEISVRGVKIGVADDSQEAAGKNCSH 240

QY 638 FFOLKNISFCGYHPKNNKYFGFITKHPADHRFACHVVFVSEDSKALAESVGRAPQOQFYKQ 697
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 FFOLKNISFCGYHPKNNKYFGFITKHPADHRFACHVVFVSEDSKALAESVGRAPQOQFYKQ 300

QY 698 FVEYTCPTEDIYLE 711
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 FVEYTCPTEDIYLE 314

RESULT 14
US-08-859-201-4
; Sequence 4, Application US/08859201
; Patent No. 5880261
; GENERAL INFORMATION:
; APPLICANT: Waeber, G
; APPLICANT: Nicod, P
; TITLE OF INVENTION: Transcription Factor Islet-Brain 1 (IB1)
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Luann Cserr
; STREET: 750 Arimo Avenue
; CITY: Oakland
; STATE: California
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,201
; FILING DATE: 20-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9706731.8
; FILING DATE: 03-APR-1997
; APPLICATION NUMBER: GB 9709920.4
; FILING DATE: 15-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Luann Cserr
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: ME A9702
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-859-201-4
Query Match 37.6%; Score 1427; DB 2; Length 270;
Best Local Similarity 99.3%; Pred. No. 2.3e-115;
Matches 268; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 442 CLSEDSPTDPDPVHFSKKFLNVFMGSRSSSAESFGLFSCIINGEQETHRAIRFRVFP 501
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CLSEDSPTDPDPVHFSKKFLNVFMGSRSSSAESFGLFSCIINGEQETHRAIRFRVFP 60

QY 502 RHDELELEVDPLLVLAEDYWEAYNMRTGARGVFPAYIAEVTKEPEHMAALAKNS 561
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 RHDELELEVDPLLVLAEDYWEAYNMRTGARGVFPAYIAEVTKEPEHMAALAKNS 120

QY 562 DWVDQFRVFLGSVQVPYHKGNDVLCAMQKIATTTTTRRLTVHFNPSSCVLEISVRGVKIG 621
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 DWVDQFRVFLGSVQVPYHKGNDVLCAMQKIATTTTTRRLTVHFNPSSCVLEISVRGVKIG 180

QY 622 VKADDSQEAAGKNCSHFFOLKNISFCGYHPKNNKYFGFITKHPADHRFACHVVFVSEDSK 681
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 VKADDSQEAAGKNCSHFFOLKNISFCGYHPKNNKYFGFITKHPADHRFACHVVFVSEDSK 240
```

0" 423 VESA YCEEYEEA DDDDDACI SEDSTDDDDVHESKKEI NVEMSCPDSSSSESCIESC 182

556	GATLLGSD-----QVP----	EDASPDSPDLTFSKKEFLNFVFNSTSRSSSTSEFGLFSC	604
483	IIINGEEQEOTHRAIFRVPVRHEDELELEVEDDPLLVLOQADYWEAYNMTGARGVFPAY	542	
605	VXNGEEREQTHRAVFRIPIRHPDELDELDDPVLVEAEDDFWFRGNMTGRRGVFPAPF	564	
543	YALEVTEPEPEHMAALAKNDSNDWDQPRVFXELGVSQVYPYHKGNDVLCAAMOKIATIRLTVH	602	
665	YAHAVPGPAKDLGLSKRSPCWDVDFDVLGSEVPEVCHOGNIIICAAMQKIATARKLTVH	724	
603	FNPPSSCVLEITSVRGVKIGVKADDSEAKGNCKSHFFOLKNIISFCGHYHPKNKXYFGEITK	662	
725	LRPASCDLEISRGVYKLSLGGGPE---FQRCSHFFQMKNIISFCGHPRNSCYFGEITK	781	
663	HPADHRFACHVFVSEDSKALAESVGRAPQOYKQVFVETCPTEDIYLE	711	
782	HPLLSRACHVFVYSOESMRVARSVGRAFLEYIOHHLAFACPTEDIYLE	830	

